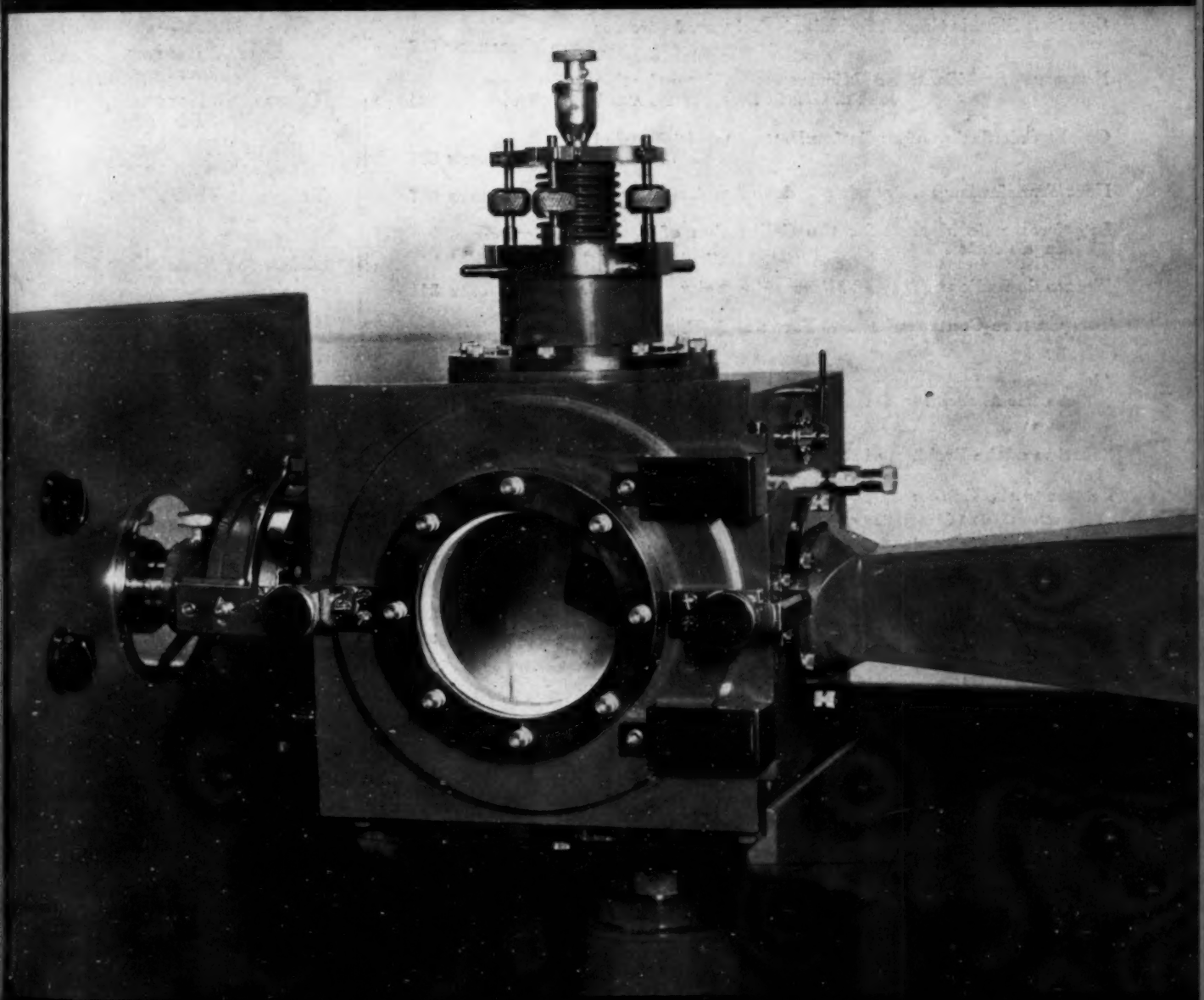
JOURNAL OF APPLIED PHYSICS



JOURNAL OF APPLIED PHYSICS

Vol. 19, No. 9 SEPTEMBER, 1948

In This Issue

		Page
Decay of Shock Waves	B. Cassen and J. Stanton	803
Surface Currents on a Conducting Sphere Excited by a Dipole	C. H. Papas and Ronold King	808
Spurious Signals Caused by Noise in Triggered Circuits	David Middleton	817
Frequency Stabilization of Microwave Oscillators by Spectrum Lines. II.	José L. García De Quevedo and William V. Smith	831
On the Calculation of Radiation Patterns of Dielectric Rods	R. B. Watson and C. W. Horton	836
Heat Transfer from a Rotating Disk to Ambient Air	Carl Wagner	837
Interferometric Method for the Calibration of Electron Microscope Magnification	J. L. Farrant and A. J. Hodge	840
Electronic Radiography and Microradiography	Jean-Jacques Trillat	844
Temperature-Controlled Frevel Focusing X-Ray Powder Camera	V. Vand	852
Absorption of Incident Quanta by Atoms as Defined by the Mass Photoelectric Absorption Coefficient and the Mass Scattering Coefficient	John A. Victoreen	855
Solution of the Problem of Rapid Scanning for Radar Antennae	R. F. Rinehart	860
Letters to the Editor:		
Electron Optical Observation of Magnetic Fields	L. Marton	863
Observation of Corrosion Cracking without Stress	W. D. Robertson and H. H. Uhlig	864

In Every Issue

Calendar of Meetings	830
New Booklets	865
Here and There	866
Index to Advertisers	xxiii

BOARD OF EDITORS

Elmer Hutchisson, *Editor*

Associate Editors

R. Bowling Barnes
Otto Beeck
W. F. Busse
W. James Lyons
Philip Morrison
Paul E. Sabine
Frederick Seitz
J. R. Townsend
John Von Neumann

Publication of the American Institute of Physics

ADMINISTRATIVE STAFF

Henry A. Barton, *Director*
Cleveland Norcross, *Assistant Director*
Ruth F. Bryans, *Publication Manager*
Theodore Vorburger, *Business Manager*
John T. Tate, *Adviser on Publications*

GOVERNING BOARD

George R. Harrison, *Chairman*
Wallace Waterfall, *Secretary*
George B. Pegram, *Treasurer*

Luis W. Alvarez
Robert B. Brode
E. U. Condon
A. J. Dempster
W. F. Fair, Jr.
Harvey Fletcher
R. C. Gibbs
W. V. Houston
Paul E. Klopsteg
F. W. Loomis
Philip M. Morse
A. H. Pfund
John T. Tate
M. W. White
A. G. Worthing

The Journal of Applied Physics, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as Physics. The Journal of Applied Physics publishes editorials, news of physicists, and reviews, as well as technical papers of applied physics.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio.

Subscriptions and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pa. or to the American Institute of Physics, 57 East 55 Street, New York 22, N. Y.

Subscription Price

	U.S. and Canada	Elsewhere
To members and associate members of the American Institute of Physics	$7.00	$ 8.00
To all others	9.00	10.00

Back Number Prices

Complete set: Vol. 1, July, 1931–Vol. 18, 1947—$132.00
Yearly back number rate: $10.00
Single copies: $1.00 each.

Changes of address should be sent to the Business Manager, American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics.

Advertising rates supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879.
Accepted for mailing at the special rate of postage provided for in the Act of February 28, 1925, authorized May 2, 1932.

Copyright 1948, by the American Institute of Physics.

Journal
of
Applied Physics

Volume 19, Number 9 September, 1948

The Decay of Shock Waves

B. Cassen and J. Stanton
Pasadena Physics Section, Naval Ordnance Test Station, Pasadena, California
(Received July 2, 1947)

The instantaneous rate of decay or intensification of a shock front in any fluid for which dissipative effects are localized in a shock zone is shown to be directly related to the coexisting pressure gradient in the wake of the wave adjacent to the front. The higher time derivatives of the shock velocity are also explicitly related, although in a more complicated fashion, to the higher spatial derivatives of wake pressure. Thus the experimental observation of the velocity or some related variable of a blast or ballistic shock wave can be used to ascertain properties of the fluid motion in the wake of the shock. Conversely, a knowledge of the wake profile at any instant can be employed to determine the corresponding rate of decay of the shock front. The general theory is applicable to shocks of all strengths where an equation of state can be applied graphically or analytically.

INTRODUCTION

IT is commonly known that blast waves decay with time, not only on account of spatials divergence, but also because of the dissipative processes occurring within the shock front. Free energy is flowing toward the shock discontiniuty at a rate determined by the local pressure, particle velocity, and the gradients of these quantities in the immediate wake. Ordinarily the rate of free energy degradation at the front, plus the rate of weakening caused by divergence, exceed the forward influx of free energy and the wave decays. But intensification may occur if the influx predominates. This latter event is rarely observed in blast waves. However, photographic evidence has been obtained in two instances of an intense blast which, emerging with strong spherical divergence from a small opening in a container, maintained nearly constant shock velocity for several hundred microseconds. During this interval, reinforcement from the rear compensated the divergence weakening although it was insufficient to cause intensification of the shock.

THEORY OF FIRST DERIVATIVE RELATIONSHIPS

The theory of the relationship of the rate of change of quantities pertaining to a shock front to the fluid motion in the wake is general and not difficult to develop. Relationships between second and higher derivatives are more involved, although still obtainable by straightforward algebra. The general theory of the first derivatives can be developed simultaneously for plane, cylindrical, and spherical shock waves by use of an index n, which is equal to 0 for the plane case, 1 for the cylindrical, and 2 for the spherical. The theory also applies to plane oblique shocks; such shocks, as is well known, can be treated as normal shocks subject to a tangential velocity transformation.

FIG. 1. Wake pressure distribution to produce decay or intensification of plane waves.

The equations of motion and continuity are written in Lagrange form as follows:

$$\begin{aligned} \partial v/\partial t &= -(x/z)^n(\partial P/\partial z); \\ V &= (x/z)^n(\partial x/\partial z). \end{aligned} \tag{1}$$

Here v = particle velocity, x = spatial position or radius, $P = p/\rho_0$ = pressure/initial density, $V = \rho_0/\rho$ = initial density/density, z = Lagrange label of particles, and t = time. The adiabatic equation of state which obtains in the wake of the shock is expressed quite generally. If S is the entropy, let this be written as:

$$V = V(S, P). \tag{2}$$

Time differentiation of the equation of continuity gives

$$\begin{aligned} (x/z)^n(\partial v/\partial z) + (nVv/x) &= \partial V/\partial t \\ &= (\partial V/\partial P)(\partial P/\partial t) + (\partial V/\partial S)(\partial S/\partial t). \end{aligned} \tag{3}$$

It is assumed that in the wake of the shock $\partial S/\partial t = 0$, but not necessarily $\partial S/\partial z$. It follows that the equations determining fluid motion in the wake are reduced to

$$\begin{aligned} \partial v/\partial t &= -(x/z)^n(\partial P/\partial z); \\ (x/z)^n(\partial v/\partial z) + (nVv/x) &= (\partial V/\partial P)(\partial P/\partial t). \end{aligned} \tag{4}$$

These equations are most conveniently considered in a coordinate system at rest with respect to the shock front in Lagrange labeled space, i.e., one for which the particle labeled z is replaced by $Z \equiv z - z_1$ and $T \equiv t$. Here $z_1(t)$ is the Lagrange position of the shock front, and $U \equiv dz_1/dt$ is the shock velocity.

Then $\partial/\partial z = \partial/\partial Z$ (represented by dash superscript in what follows) and $\partial/\partial t + U(\partial/\partial z) = \partial/\partial T$ (represented by dot derivative in what follows). After transformation, Eqs. (4) become

$$\begin{aligned} \dot{v} - Uv' &= -(x/z)^n P'; \\ (x/z)^n v' + nVv/x &= \partial V/\partial P(\dot{P} - UP'). \end{aligned} \tag{5}$$

These two equations are linear in v' and P' and

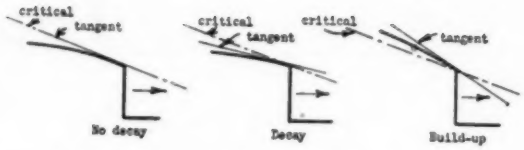


FIG. 2. Wake pressure distribution to produce decay or intensification of diverging waves.

can be solved explicitly for these quantities as

$$\begin{aligned} P' &= \left[-\dot{v} + U\left(\frac{z}{x}\right)^n \left(\frac{\partial V}{\partial P}\dot{P} - \frac{nVv}{x}\right)\right] \Big/ \\ &\qquad \left[\left(\frac{x}{z}\right)^n + U^2\frac{\partial V}{\partial P}\left(\frac{z}{x}\right)^n\right], \\ v' &= \left[\frac{\partial V}{\partial P}\dot{P} - \frac{nVv}{x} + U\frac{\partial V}{\partial P}\left(\frac{z}{x}\right)^n \dot{v}\right] \Big/ \\ &\qquad \left[\left(\frac{x}{z}\right)^n + U^2\frac{\partial V}{\partial P}\left(\frac{z}{x}\right)^n\right]. \end{aligned} \tag{6}$$

When specialized to particular values at the shock front, where $x = z = z_1$, one obtains

$$\begin{aligned} P_1' &= \left[-\dot{U}(v_U - URP_U) - \frac{nV_1v_1U}{z_1}\right] \Big/ [1 + RU^2], \\ v_1' &= \left[\dot{U}(RP_U - URv_U) - \frac{nV_1v_1}{z_1}\right] \Big/ [1 + RU^2], \end{aligned} \tag{7}$$

where $R \equiv (\partial V/\partial P)_{Z=0}$ and U used as a subscript indicates differentiating with respect to U at the shock front. As will be shown, all quantities appearing on the right-hand side of these expressions can be determined in terms of U and $\dot{U}$, when the equation of state of the fluid is given.

For a steady plane wave $n = 0$, and $\dot{U} = 0$, so that $P' = 0$ and the tangent to the pressure distribution in Lagrange space is horizontal at the shock front. If the pressure falls off behind the front with a tangent of positive slope, the shock will decay, and conversely a negative slope indicates a shock that will build up (see Fig. 1).

For a diverging wave, at a given value of z_1 from the origin or axis, there will be a critical value of gradient, depending on the strength of the shock and the nature of the fluid, at which the free energy influx will balance the divergence loss. For gradients smaller than this critical

value the wave will decay; for steeper gradients it will be intensified (see Fig. 2).

The nearer the shock front is to its source, the steeper will be its critical gradient. The critical gradient is obtained by setting $\dot{U}=0$ in (7), giving

$$P'_{\text{critical}}=-\frac{nV_1v_1U}{(1+RU^2)(z_1)}.$$

These considerations apply equally well to the formation of a tail or trailing shock wave such as those observed in the case of ballistic N-waves.* A pressure gradient in excess of the critical must be produced in the wake of the projectile as a necessary preliminary to the development and growth of the shock. As the critical gradients are still steeper in the case of spherically diverging waves it appears very unlikely, but not impossible under unusual conditions, that a blast wave will ever develop a tail shock. Also, from this point of view, a projectile with a long tapering tail will probably not give a ballistic wave with a tail shock.

The relationships between wake gradients and time derivatives expressed by (7) can easily be specialized for an ideal gas. The Rankine-Hugoniot relations relating the shock variables can be expressed analytically in this case for which the adiabatic equation of state is $P=A(S)V^{-\gamma}$, in the well-known form

$$\begin{aligned}P_1&=C^2[2\gamma(U/C)^2+(1-\gamma)]/(\gamma(1+\gamma)),\\ v_1&=2[U-C^2/U]/(\gamma+1),\\ V_1&=[(\gamma-1)+2(C^2/U^2)]/(\gamma+1),\\ R&=-V/\gamma P\end{aligned}\tag{8}$$

(where C is the normal velocity of sound) and corresponding time derivatives. This relation between the shock variables makes possible the explicit expression of P' and v' in terms of U and $\dot{U}$:

$$P_1'=\frac{-2\dot{U}}{(\gamma+1)^2(U^2-C^2)}\{[1+C^2/U^2][2\gamma U^2+(1-\gamma)C^2]+2[U^2(\gamma-1)+2C^2]\}$$
$$-\frac{2n}{z_1(\gamma+1)^3U^2}\{[(\gamma-1)U^2+2C^2][2\gamma U^2+(1-\gamma)C^2]\},\tag{9}$$

and

$$v_1'=-\frac{2\dot{U}[3+C^2/U^2][(\gamma-1)U^2+2C^2]}{(\gamma+1)^2U(U^2-C^2)}-\frac{2n}{z_1U^3(\gamma+1)^3}\{[(\gamma-1)U^2+2C^2][2\gamma U^2+(1-\gamma)C^2]\}.\tag{10}$$

Some applications of these expressions will be considered after a brief discussion of the general relationships connecting higher derivatives.

THEORY OF HIGHER DERIVATIVES

The relationship for second and higher derivatives may be obtained directly from (6) by differentiation and substitution but unfortunately involves a large number of terms, although only in simple algebraic combinations. The dot derivatives should be taken first, i.e., $\dot{P}'$ and $\dot{v}'$, since the evaluation of these is required for the later calculation of P'' and v''. Recalling that

$$\begin{aligned}&\dot{z}=U,\quad \dot{x}=\partial x/\partial t+U(\partial x/\partial z)=v+UV(z/x)^n,\\ &\dot{R}=(\partial R/\partial P)\dot{P}+(\partial R/\partial S)\dot{S},\quad S'=\dot{S}/U,\\ &R'=(\partial R/\partial P)P'+(\dot{S}/U)(\partial R/\partial S),\\ &V'=RP'+(\dot{S}/U)(\partial V/\partial S),\end{aligned}$$

the process is straightforward and is not shown explicitly. The result P_1'' for P'' at the shock front can be shown to be

$$\begin{aligned}P_1''=&-(U^2R+1)^{-2}\{(nP'/z_1)(3U^2R+1)(V-1)\\ &-2UR(-(nv/z_1)P'+\dot{U}v'-(\partial^2v/\partial T^2)\\ &+2U^3R[n/z_1(2v'V-v'+vRP'\\ &+v(\dot{S}/U)(\partial V/\partial S)-(V^2v/z_1))\\ &+(UP'-\dot{P})((\partial R/\partial P)P'+(\dot{S}/U)(\partial R/\partial S))]\\ &+(U^2R-1)[RP'\dot{U}-R\ddot{P}-(\partial R/\partial P)(\dot{P}-UP')^2\\ &+(n/z_1)(vv'+V\dot{v}-UVv'-Vv^2/z_1\\ &+vR\dot{P}-vRUP')]\},\end{aligned}\tag{11}$$

where all quantities on the right are evaluated at the shock front. In principle then, a knowledge of U, $\dot{U}$, $\ddot{U}$,—along with the equation of state and derived Hugoniot equation of state—would enable a Taylor's series to be constructed for P or v as a series in Z and T with coefficients functions of T.

* DuMond, Cohen, Panofsky, and Deeds, J. Acous. Soc. Am. **18**, 97 (1946).

An interesting special case involving higher order gradients is that of the constant velocity spherically divergent wave found experimentally as previously mentioned. In this case $\dot{U}$, $\ddot{U}$, vanish and in the wake of the shock S is independent of z as well as t. Then

$$\begin{aligned}P_1' &= nV_1v_1U/(z_1(1+RU^2)) = f(U)/z_1,\\ P_1'' &= -(U^2R+1)^{-2}\{(nP'/z_1)[(V_1-1)(3U^2R+1)\\ &\quad +v_1UR(U^2R+3)]+(nv_1'/z_1)[U(3V_1U^2R\\ &\quad +V_1-2U^2R)+v_1(U^2R-1)]\\ &\quad +(nvV/z_1^2)[-U^2R(2U(\partial^2v/\partial V_1)+v_1)+v_1]\\ &\quad +U^2P_1''^2(\partial R/\partial P)(U^2R+1)\}\\ &= g(U)/z_1^2. \qquad (12)\end{aligned}$$

This is of form $g(U)/z_1^2$, since P_1' and v_1' are of the form $f(U)/z_1$. Similarly, $P_1''' = h(U)/z_1^3$. Thus

$$\begin{aligned}P &= P_1+P_1'Z+P_1''Z^2/2!+\cdots\\ &= P_1+(f(U)/U)(Z/T)+(h(U)/2U^2)(Z/T)^2\\ &\quad +\cdots = P(Z/T). \qquad (13)\end{aligned}$$

At the origin, $Z=-UT$, so that the pressure there is always constant and determined only by U. From dimensional considerations an explosion with constant pressure at the origin must produce a distribution which is a function of Z/T.** For such an explosion the pressure distribution might look like that indicated in Fig. 3 and be expected to be found where a high pressure is suddenly produced in a large container with a small opening so that the source of pressure is maintained for a considerable length of time. This situation is in contrast to the opposite extreme in which the energy release occurs in a negligibly small time interval. In most explosions energy is released over a small but finite time interval, and this generation of energy is spread over a finite region of space. For such a finite source of given explosive energy release, the more the concentration of the explosion in space and time the greater will be the strength of the initial shock but the faster will be its initial rate of decay. If the pressure is maintained for a longer time by slower release of energy, the initial shock will be less intense but also will decay less rapidly since it is augmented at a later time by the influx of free energy. However, at long enough times compared with the temporal spread of the explosion and at great enough distances compared with the region of energy release, the shock strength should approach a value determinable from the total energy of the source and the distance from the origin.

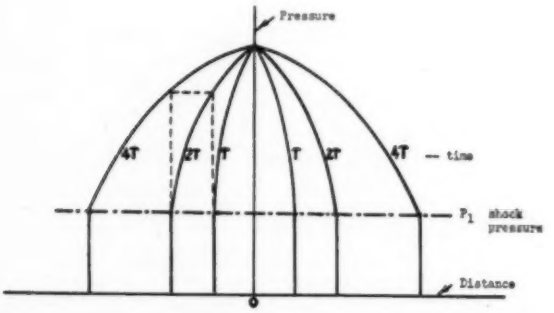


FIG. 3. Propagation, at constant velocity, of a shock originating from an explosion at constant pressure.

** This follows as pressure of dimensions $(ML^{-1}T^{-2})$ and the undisturbed air density of dimensions (ML^{-3}) are the only relevant variables. Their ratio of dimensions (L^2T^{-2}) is independent of mass and is the only combination of L and T determined by the data.

APPLICATION TO WEAK SHOCKS

It is of considerable interest to observe the forms that the wake pressure derivatives take when the Mach number approaches unity. Equation (9) reduces to the form

$$\begin{aligned}P_1' &= -2(C^2/\gamma+1)[(4\dot{U}/U^2-C^2)+n/z_1]\\ &= -2(C/\gamma+1)[(2\dot{v}_1/v_1)+(nC/z_1)]. \qquad (14)\end{aligned}$$

Similarly, Eq. (12) reduces to

$$P_1'' = (4C/\gamma+1)[3(\partial^2v/\partial T^2)_1/v_1^2+4\dot{v}_1^2/v_1^3]. \qquad (15)$$

In the plane case, $n=0$, the differential equations (4) can be solved for weak waves where the entropy is practically constant. For this case

$$\partial v/\partial t = -\partial P/\partial z; \quad \partial v/\partial z = R(\partial P/\partial t). \qquad (16)$$

Eliminating v through differentiation gives

$$-(\partial^2P/\partial z^2) = R(\partial^2P/\partial t^2)+(\partial R/\partial P)(\partial P/\partial t)^2. \qquad (17)$$

The equation of state reduces to

$$V = [P\gamma/C^2]^{-1/\gamma},$$

where $A(S)$ is considered constant. Also,

$$R = \partial V/\partial P = -V/\gamma P; \qquad (18)$$

whence,

$$\partial R/\partial P = (1+\gamma/\gamma^2)(V/P^2).$$

Then (17) becomes

$$\begin{aligned}&-\gamma A^{-1/\gamma}P^{1+1/\gamma}\partial^2P/\partial z^2\\ &\quad = \partial^2P/\partial t^2+(1+\gamma/\gamma)\cdot(1/P)(\partial P/\partial t)^2. \qquad (19)\end{aligned}$$

A time-dependent solution of this equation is

$$P=p/\rho_0=(p_0/\rho_0)(z/ct)^{2/1+\gamma}, \quad (20)$$

giving

$$v=(2c/\gamma-1)[(p/p_0)^{(\gamma-1)/2\gamma}-1], \quad (21)$$

and

$$\partial p/\partial z=(2\gamma p_0/1+\gamma)(1/ct)(z/ct)^{(\gamma-1)/\gamma+1}, \quad (22)$$

if $p=p_0(1+\delta)$, where δ, the over-pressure in atmospheres, is small compared with unity, then

$$v/c=[2/(\gamma-1)][1+((\gamma-1)/2\gamma)\delta+\cdots-1] =\delta/\gamma. \quad (23)$$

But this is also the approximate Hugoniot condition between particle velocity and over-pressure for very weak shocks. Therefore, the abolve solution (20) for pressure can be cut off by a shock boundary at any value of z where δ is small enough. As a case in point, Fig. 4 depicts how this time-dependent solution to the hydrodynamic equations can be terminated by head and tail shocks, H and T, to represent a plane N-wave. The pressure distribution is not a straight line between the terminating shocks.

When $z=ct$, $p=p_0$ so that the plane of pressure p_0 travels with constant velocity c in Lagrange space and therefore also in actual space for weak disturbances.

The pressure gradient at the terminating shocks will differ only very slightly from the slope of the pressure at $p=p_0$ so that the rate of decay of the shock terminations can be closely determined by putting this value of gradient in (14), giving

$$(2\gamma/1+\gamma)(p_0\rho_0/Ct)=(-4C/\gamma+1)(\dot{v}_1/v_1), \quad (24)$$

whence

$$\dot{v}_1/v_1=-1/2t \quad \text{or} \quad \delta_1/\rho_1 C=v_1=kt^{-\frac{1}{2}}, \quad (25)$$

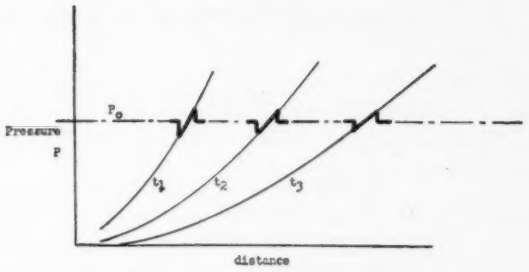


FIG. 4. Progress in time for a plane N-wave constructed from a time-dependent solution of the hydrodynamic equations terminated by weak shocks.

where k is a constant. Therefore, to this degree of approximation both the head and tail shocks should decay inversely as the square root of the time. To do this they must move away from the plane $z=Ct$, at which $p=p_0$, as time progresses. Assuming the length λ of the N-wave is $\lambda=z_1-Ct$, and recalling that

$$(1+\delta_1)(1+\gamma)/2\gamma=1+(1+\gamma/\gamma)\delta_1 =z_1/Ct=(\lambda+Ct)/Ct=1+\lambda/Ct, \quad (26)$$

it follows that

$$\lambda=(1+\gamma/\gamma)Ct\delta_1.$$

Therefore, from (25)

$$\lambda=(1+\gamma/\gamma)Ct\cdot k\rho_0 Ct^{-\frac{1}{2}}, \quad \sim t^{\frac{1}{2}}. \quad (27)$$

The length of the plane N-wave increases as the square root of the time. It should be noted that these laws of decay of weak plane waves are independent of the γ of the fluid.

Because of the explicit occurrence of x in Eqs. (1) for cylindrically and spherically divergent waves, the method developed here for the study of the decay of plane waves is not directly applicable. Further studies are being made to determine whether as rigorous a method can be developed for diverging waves.

Surface Currents on a Conducting Sphere Excited by a Dipole*

C. H. Papas and Ronold King
Cruft Laboratory, Harvard University, Cambridge, Massachusetts
(Received January 23, 1948)

This paper treats the problem of determining the current distribution on the surface of a perfectly conducting sphere when driven by a dipole antenna erected on its surface. Curves of the real and imaginary parts of the surface currents are given for the case of a half-wave dipole and various radii of the sphere.

I. INTRODUCTION

THE distribution of currents on the surface of a metal spheroid excited by an adjacent antenna or by a cavity is of great practical interest (see Fig. 1). The analytical problem involved is closely related to the determination of the diffracted field of a plane wave incident upon a sphere or spheroid,[1] which has been studied by many investigators. The rigorous solution of the general case in which the wavelength of the incident plane wave is of the same order of magnitude as the dimension of the diffracting sphere is due to L. Brillouin. In practice a plane wave is approximated by the field of a radiator situated a great many wavelengths from the diffracting sphere or prolate spheroid. This field induces surface currents which, in turn, set up an electromagnetic field. This induced field is so disposed that on the surface of the sphere or spheroid (assumed perfectly conducting) the sum of the tangential components of the incident and induced E-fields vanishes and the sum of the tangential components of the H-fields equals the surface current density. Since the scalar wave equation is separable in spherical (and spheroidal) coordinates, these boundary conditions permit the determination of the unknown constants of the infinite series solution for the induced E- and H-fields.

If the distant radiator is brought close to the diffracting body, the incident field is no longer a plane wave and the analytical problem becomes so complicated that the general procedure described above has to be abandoned. An alternative procedure has been formulated by Feld[2] who has applied Lorentz' lemma to a linear antenna over a perfectly conducting sphere. The current in the antenna is assumed to be sinusoidal; the coefficients of the modal currents on the sphere are determined by the well-known Fourier-Lamé method.[3]

By applying the general reciprocity theorem (Appendix A) to a closed metallic surface with a linear antenna erected upon it, an integral equation is obtained. In order to solve this integral equation by the Fourier-Lamé method it is necessary to restrict the closed metallic surface to a shape that may be described simply using coordinates in which the homogeneous scalar wave equation,

$$\nabla^2\psi + k^2\psi = 0,$$

is separable.[4] Another limitation (which is of purely practical importance) is the availability of numerical tables of the functions used in the solutions. For example, if the closed metallic surface is a sphere, the "natural" coordinates are spherical which permit separation of variables and give as angular solutions the Legendre polynomials and as radial solutions the spherical Hankel functions. Both the Legendre polynomials[5] and spherical Hankel functions[6] have been tabulated. Similarly, if the metallic surface is a prolate spheroid, the prolate spheroidal coordinates[1] ξ, η, ϕ are used. These also permit

* The research reported in this document was made possible through support extended Cruft Laboratory, Harvard University, jointly by the Navy Department (Office of Naval Research) and the Signal Corps, U. S. Army under Contract N5 ori-76, T. O. I.

[1] J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill Book Company, Inc., New York, 1941).

[2] J. N. Feld, Doklady U.S.S.R. **51**, No. 3, 203–206 (1946).

[3] G. Grunberg, J. Phys. U.S.S.R. **10**, No. 4, 301–320 (1946).

[4] L. P. Eisenhart, Annals Math. **35**, No. 2, 284–305 (1934).

[5] *Tables of Associated Legendre Functions* (Columbia University Press, New York, 1945).

[6] *Tables of Spherical Bessel Functions* (Columbia University Press, New York, 1945).

separation of variables and yield the functions $Se_l^1(\eta)$ as "angular" solutions and $Re_l^1(\xi)$ as "radial" solutions. These functions have been tabulated.[7] It is evidently possible to solve for the surface currents on spheres, prolate spheroids, and oblate spheroids, and from these to determine the transmitting and receiving characteristics of the array consisting of linear antenna and closed surface.

Although the spheroidal functions have been tabulated, the additional computation necessary to determine the surface currents is considerable. Therefore, the simpler problem of a linear antenna over a sphere has been carried through first.

II. LINEAR ANTENNA OVER SPHERE

For an array consisting of a linear antenna of height h over a perfectly conducting sphere of radius a, the total current crossing a parallel of latitude θ on the sphere is (Appendix C)

$$I_{1S}(\theta)=\sum_{n=0}^{\infty} B_n P_n(\cos\theta), \tag{1}$$

where

$$\begin{aligned} B_n=(I_{\max}/2)[&[\rho_{n-1}(kd)/\rho_{n-1}'(ka)] \\ &-\cos kh[\rho_{n-1}(ka)/\rho_{n-1}'(ka)] \\ &-[\rho_{n+1}(kd)/\rho_{n+1}'(ka)] \\ &+\cos kh[\rho_{n+1}(ka)/\rho_{n+1}'(ka)]], \end{aligned} \tag{2}$$

and $k=2\pi/\lambda$, a is the radius of sphere, h is the height of antenna, $d=a+h$. It was assumed in the derivation of (1) that the current along the antenna has the form

$$I_{1A}(r)=I_{\max}\sin k(d-r),$$

and that this distribution is independent of currents on the sphere. For very thin antennas driven at the junction with the sphere a sinusoidal distribution is a good approximation for lengths differing considerably from $\lambda/2$. In general, a sinusoidal current cannot be maintained by a single generator but requires a distribution of generators along the entire antenna.

When the length of the antenna is $\lambda/4$, the current is maximum, $I_{\max}$, at the base ($r=a$) and zero at the top ($r=a+h$). Since $kh=\pi/2$,

[7] Stratton, Morse, Chu, and Hutner, *Elliptic Cylinder and Spheroidal Wave Functions* (Massachusetts Institute of Technology Press, Cambridge, 1941).

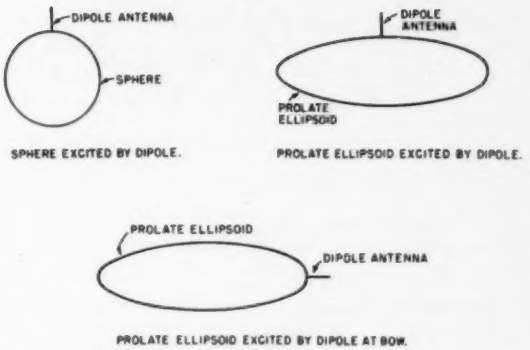


FIG. 1. Some closed metallic surfaces excited by dipoles which may be solved for surface currents by the method of this paper.

the expression for B_n reduces to

$$\begin{aligned} B_n=(I_{\max}/2)[&[\rho_{n-1}(kd)/\rho_{n-1}'(ka)] \\ &-[\rho_{n+1}(kd)/\rho_{n+1}'(ka)]]. \end{aligned} \tag{3}$$

Calculations have been made for various values of the radius a with $kh=\pi/2$ using (3). In order to compute terms such as the ratio

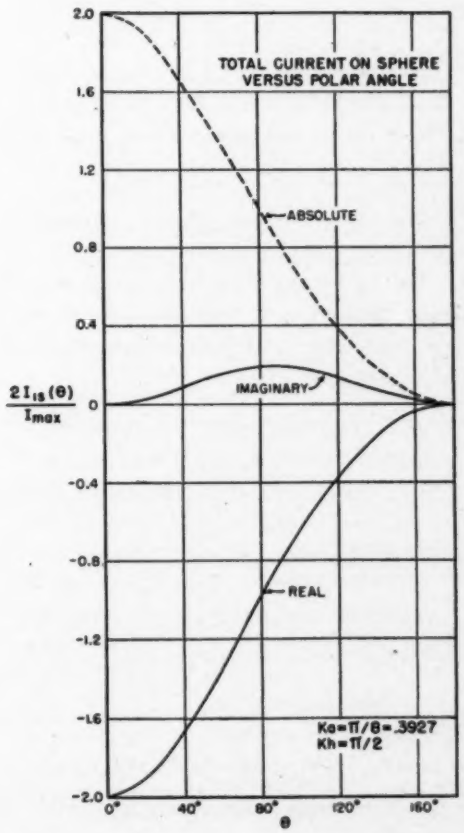


FIG. 2. Total current on sphere *versus* polar angle for $ka=\pi/8$.

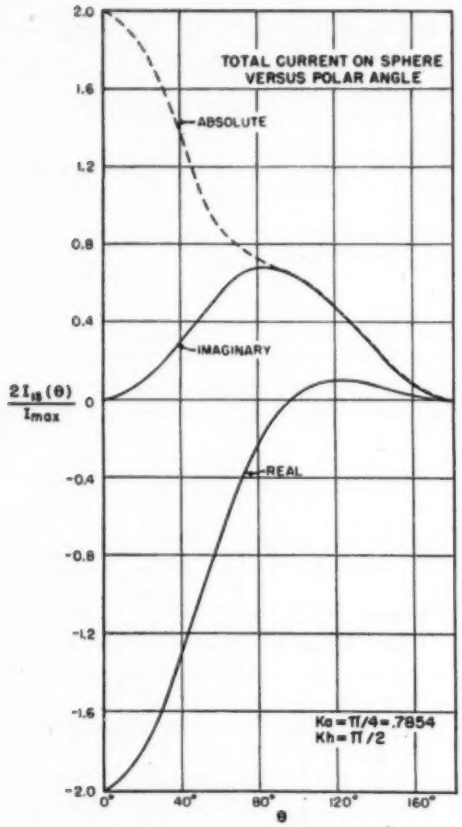


FIG. 3. Total current on sphere *versus* polar angle for $ka=\pi/4$.

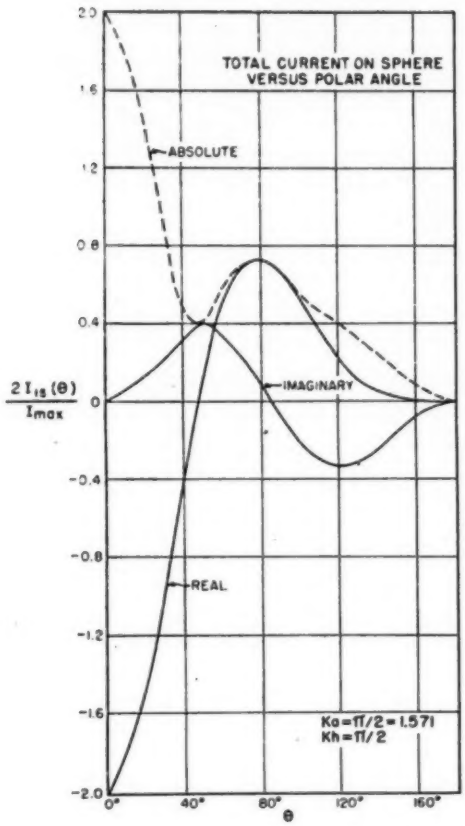


FIG. 4. Total current on sphere *versus* polar angle for $ka=\pi/2$.

$\rho_{n-1}(kd)/\rho_{n-1}'(ka)$ it was found convenient to follow a computational procedure used by L. Infeld.[8] By factoring this ratio it is seen that one of the factors is a function which satisfies Riccati's equation and for which there is a recurrence relation. That is,

$$\rho_{n-1}(kd)/\rho_{n-1}'(ka)=\rho_{n-1}(ka)/\rho_{n-1}'(ka)\cdot\rho_{n-1}(ka+kh)/\rho_{n-1}(ka)=\zeta_{n-1}(ka)\cdot\rho_{n-1}(ka+kh)/\rho_{n-1}(ka), \quad (4)$$

where

$$\zeta_{n-1}(ka)=\rho_{n-1}(ka)/\rho_{n-1}'(ka).$$

This zeta satisfies Riccati's equation and equals the zeta used by Infeld with the n equal to his $2m+1$. Letting $x=ka$, the recurrence formula is

$$\zeta_n(x)=[nx\zeta_{n-1}(x)-x^2]/[(x^2-n^2)\zeta_{n-1}(x)+nx], \quad (5)$$

and $\zeta_{-1}(x)=+i$. Starting with $\zeta_{-1}(x)=+i$, the entire set of functions $\zeta_0(x)$, $\zeta_1(x)$, $\cdots$ was obtained.

The other factor which is of the form $\rho_{n-1}(ka+kh)/\rho_{n-1}(ka)$ was computed from the fundamental definition of the functions:

$$\frac{\rho_{n-1}(ka+kh)}{\rho_{n-1}(ka)}=\frac{(x+kh)[(\pi/2(x+kh))^{\frac{1}{2}}J_{n+\frac{1}{2}}(x+kh)+(i/\cos n\pi)(\pi/2(x+kh))^{\frac{1}{2}}J_{-(n+\frac{1}{2})}(x+kh)]}{x[(\pi/2x)^{\frac{1}{2}}J_{n+\frac{1}{2}}(x)+i(1/\cos n\pi)(\pi/2x)^{\frac{1}{2}}J_{-(n+\frac{1}{2})}(x)]}. \quad (6)$$

The functions

$$(\pi/2x)^{\frac{1}{2}}J_{n+\frac{1}{2}}(x) \quad \text{and} \quad (\pi/2x)^{\frac{1}{2}}J_{-(n+\frac{1}{2})}(x)$$

are the spherical Bessel functions, tables of which are available.

[8] L. Infeld, Q. App. Math. **5**, No. 2, 113–132 (1947)

Using (6) and (5), the ratios $\rho_{n-1}(ka+kh)/\rho_{n-1}'(ka)$ were computed for $n=0, 1, 2, 3, 4\cdots 11$ when $kh=\pi/2$ and $ka=\pi/8$, $\pi/4$, $\pi/2$, $3\pi/4$, π, $5\pi/4$, $3\pi/2$, $7\pi/4$, and 2π. These ratios were substituted into (3). The values of B_n thus

obtained were substituted into (2) and $I_{1S}(\theta)$ was determined.

III. RESULTS

The real and imaginary parts of the current $I_{1S}(\theta)$ crossing the θ parallel of latitude have been computed for several values of the radius a of the perfectly conducting sphere in the special case of an antenna of length $h=\lambda/4$ erected on the sphere. Plots of these currents are shown in Figs. 2–10. The ordinate of these plots is twice the ratio of $I(\theta)$ to $I_{\max}$. Since the positive direction of current along the antenna is towards its top and the positive direction of flow on the sphere is from its North pole towards its South, it is necessary for $I_{1S}(0)$ to be opposite in sign to $I_{1A}(a)=I_{\max}$. In all the plots for $\theta=0$ the imaginary part is zero and the real part equals -2. That is,

$$2I_{1S}(0)/I_{\max}=-2 \quad \text{or} \quad I_{1S}(0)=-I_{\max}.$$

At the bottom of the sphere, $\theta=180°$, the current is zero.

In the theoretical derivation it was assumed that there was a generator or a distribution of generators somewhere in the antenna. The only restriction upon this generator is that it be independent of the angle ϕ. For example, the generator may be a single idealized slice generator or a distribution of these along the antenna. In the computed case the current distribution is assumed to be sinusoidal with a node at the top of the antenna and a loop $I_{\max}$ at its base. This requires a single slice generator to be placed at the base of a very thin antenna as indicated in Fig. 11. The current distributions shown in the plots are for just such a model.

Since an actual physical antenna of finite cross section driven from a coaxial line over an imperfectly conducting sphere necessarily differs from the ideal model assumed in the theory, actual currents must be expected to differ somewhat from the computed currents represented in the plots. However, it is safe to assume that the currents in the idealized mathematical model are

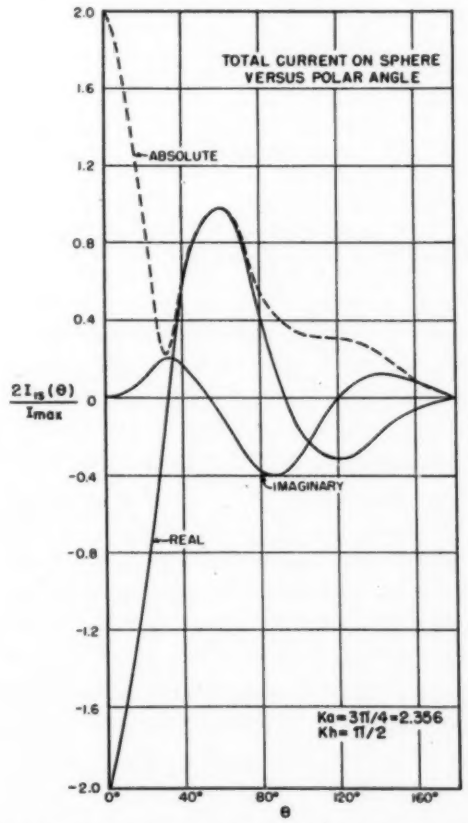


FIG. 5. Total current on sphere *versus* polar angle for $ka=3\pi/4$.

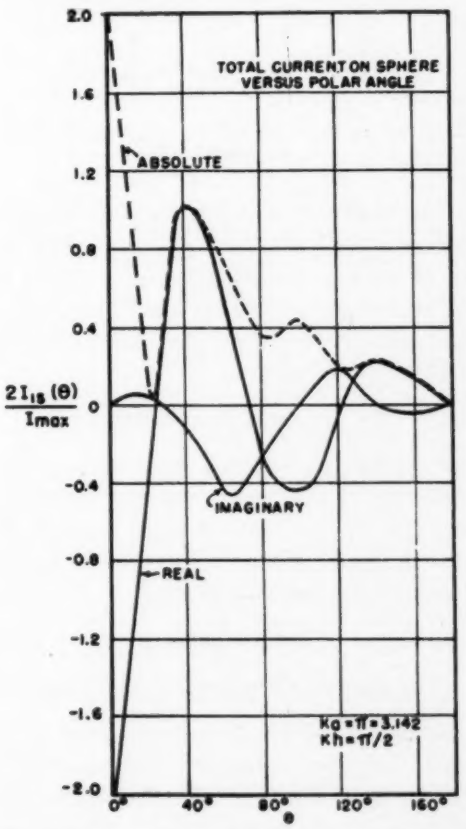


FIG. 6. Total current on sphere *versus* polar angle for $ka=\pi$.

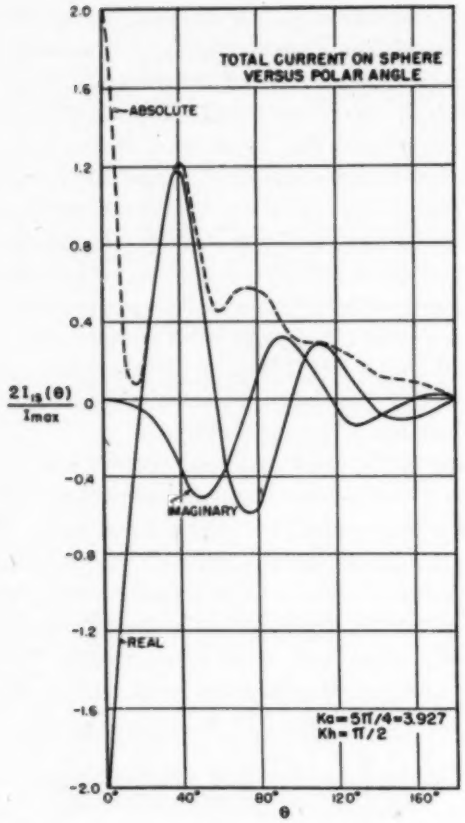


FIG. 7. Total current on sphere *versus* polar angle for $ka=5\pi/4$.

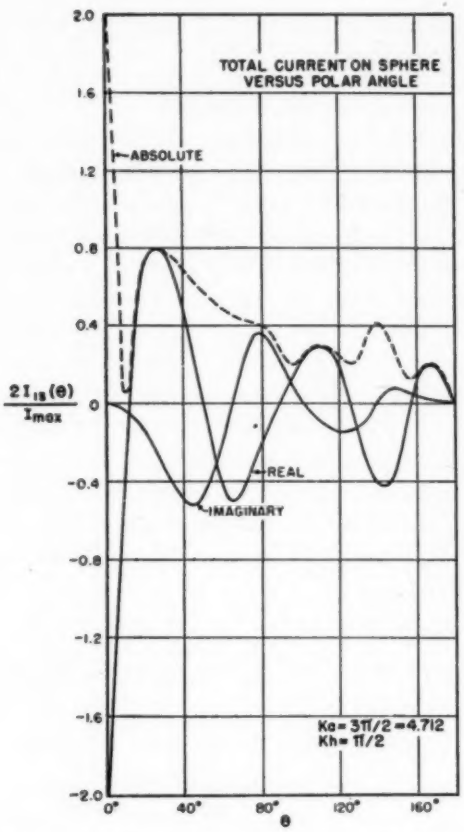


FIG. 8. Total current on sphere *versus* polar angle for $ka=3\pi/2$.

a fair approximation of actual currents in a suitably arranged physical set-up. The determination of such actual distributions using techniques developed in this laboratory for the measurement of magnitude and phase of surface currents is contemplated.

IV. ACKNOWLEDGMENTS

The writers are indebted to Dr. C. T. Tai of this Laboratory for his assistance during the preparation of this manuscript. Computations were made by Miss M. L. Richardson and Mrs. R. Stokey of the laboratory staff.

APPENDIX A

The General Reciprocity Theorem

A given source or distribution of sources maintains a current density i_1 with an associated electromagnetic field E_1, H_1. A second source maintains a current density i_2 with an associated electromagnetic field E_2, H_2. Then, according to the theorem of Lorentz,

$$\nabla\cdot(E_1\times H_2)-\nabla\cdot(E_2\times H_1)=i_1\cdot E_2-i_2\cdot E_1. \quad \text{(A1)}$$

Up to this point there is no relation either between the two sets of fields and currents or the regions in which they are defined. Let it be assumed that the two sets of fields are defined in free space and are bounded by the same geometrical surfaces. The boundary conditions on these surfaces need not be related in any other manner than that they are applied at surfaces that satisfy the same mathematical equations. Let (A1) be multiplied on each side by an element of volume $d\tau$ and integrated over a volume V in empty space that is bounded by a closed surface S and extends to infinity where all fields vanish. Since all volume densities of current vanish in empty space, the right side of (A1) when so integrated vanishes. The left side may be transformed using the divergence theorem, since the fields are continuous in the volume V.

The result is

$$\int_V \{\nabla\cdot(E_1\times H_2 - E_2\times H_1)\}d\tau$$
$$= \int_S \{\hat{n}\cdot(E_1\times H_2 - E_2\times H_1)\}d\sigma = 0, \quad \text{(A2)}$$

where $\hat{n}$ is an external normal to the volume V on the bounding surface S. The only restrictions on the two sets of fields is that they satisfy the field equations, that the volume of integration V be empty space, and that the bounding surface S have the same geometrical shape for both sets but not necessarily the same physical properties.

For periodic time dependence (A2) is valid for complex amplitudes provided both fields vary with the same frequency.

APPENDIX B

Derivation of the Integral Equation

The problem under consideration is to determine the distribution of current on the surface of a perfectly conducting sphere of radius a when excited by a given distribution of current along a linear antenna of very small cross-sectional area. The antenna is placed along the line $\theta=0$ in a system of spherical coordinates r, θ, ϕ with origin at the center of the sphere as shown in Fig. 11. Let the volume V be the entire volume in empty space outside the sphere and the antenna. Similarly, let the inner bounding surface S consist of the surface S_S of the sphere and the surface S_A of the antenna. The outer boundary is at infinity where all fields vanish.

The electromagnetic field vectors E_1 and H_1 are defined to be caused by the actual currents in the antenna and on the surface of the sphere. On the perfectly conducting sphere they must satisfy the boundary conditions:

$$\hat{n}\times E_1 = 0 \text{ on } S_S, \quad \text{(B1)}$$

$$\hat{n}\times H_1 = K_{1S} \text{ on } S_S, \quad \text{(B2)}$$

where $\hat{n}$ is an external normal to the sphere and

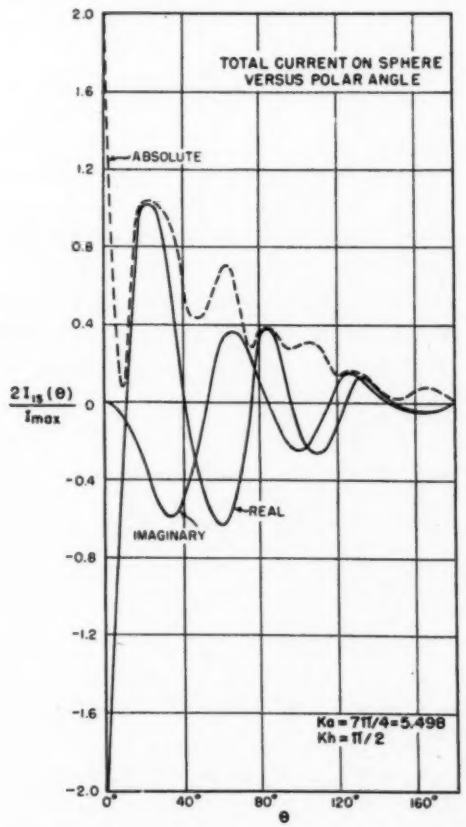


FIG. 9. Total current on sphere *versus* polar angle for $ka=7\pi/4$.

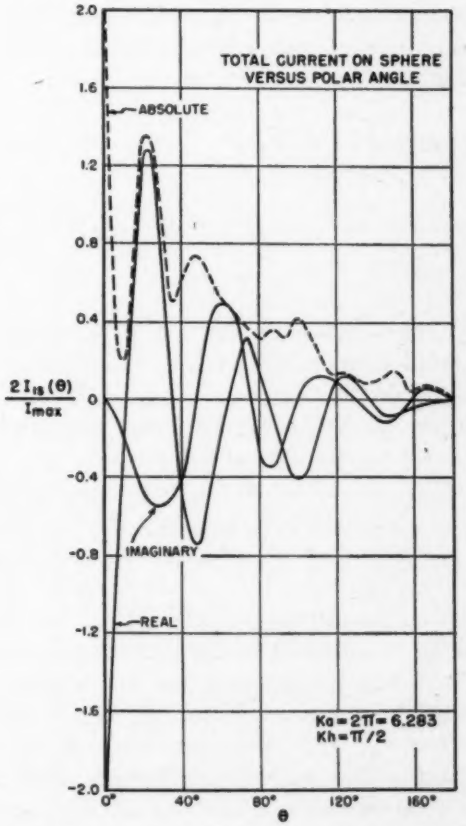


FIG. 10. Total current on sphere *versus* polar angle for $ka=2\pi$.

hence an internal normal to the volume V bounded by S, and where K_{1S} is the surface density of current on the sphere. The current i_{1A} in the antenna and K_{1S} on the sphere are maintained by an appropriate generator or distribution of generators *in the antenna* so disposed that no currents exist in the ϕ-direction (i.e., around an axis through the antenna), in the antenna or on the sphere. Hence,

$$K_{1S}=\hat{\theta}K_{1S}(\theta);$$
$$\int_0^{2\pi} K_{1S}(\theta)\cdot a\sin\theta d\phi = I_{1S}(\theta), \tag{B3}$$

where $I_{1S}(\theta)$ is the total current crossing a parallel of latitude on the sphere. Also the total current in the antenna is axial, i.e., radial in the system of spherical coordinates. Evidently, complete rotational symmetry is obtained about the axis $\theta=0$ through the antenna and the sphere. The line integral of the tangential component of the magnetic field H_1 around the surface of the antenna is equal to the total axial current $I_{1A}(r)$ in the antenna. Thus, with R_0 the radius of the antenna is

$$\oint H_1\cdot ds=\int_0^{2\pi} H_{1\phi}R_0 d\phi = \int_{\substack{\text{cross}\\ \text{section}}} i_{1A}(r)dS=I_{1A}(r). \tag{B4}$$

The electromagnetic vectors E_2 and H_2 are not actual fields related to the real currents of density i_1 in the antenna or the sphere. They constitute an auxiliary, fictitious field that is introduced to facilitate the derivation and solution of the integral equation. It has been pointed out that E_2 and H_2 must satisfy the field equations and that they must be defined on the same geometrical surfaces, i.e., on the surfaces of the sphere and the antenna, as E_1 and H_1. Similarly, i_2 and E_2 are defined in the same volume as i_1 and E_1. Let it be assumed that the volume V_A (occupied by the antenna for E_1, H_1) is empty space for E_2, H_2. Also, let the volume V_S (occupied by the perfectly conducting sphere for E_1, H_1) be constituted to contain appropriate generators or distributions of generators to maintain the following prescribed field on the surface S_S of the sphere

$$\begin{aligned} &E_{2\phi}=0; \quad E_{2\theta}=U\sin\theta P_n(\cos\theta),\\ &H_{2\theta}=0; \quad H_{2r}=0. \end{aligned} \tag{B5}$$

Thus the entire field in empty space and its boundary S_V on the sphere has the components E_{2r}, $E_{2\theta}$, $H_{2\phi}$ which must satisfy the field equations and the prescribed boundary conditions (B5).

After rearrangement of the vector products in (A2) they may be written as follows:

$$\int_{SS}\{H_2\cdot(\hat{n}\times E_1)+E_2\cdot(\hat{n}\times H_1)\} + \int_{SA}\{H_2\cdot(\hat{n}\times E_1)-H_1\cdot(\hat{n}\times E_2)\}d\sigma=0. \tag{B6}$$

Let the integrals in (B6) be considered in turn. As a result of (B1) and (B2) the first integral becomes

$$\int_{SS}\{H_2\cdot(\hat{n}\times E_1)+E_2\cdot(\hat{n}\times H_1)\}d\sigma = \int_{SS}E_2\cdot K_{1S}d\sigma=\int_{SS}E_{2\theta}K_{1S}(\theta)d\sigma = \int_0^{\pi}E_{2\theta}I_{1S}(\theta)ad\theta. \tag{B7}$$

The second and third steps follow from (B3) and (B5).

On the surface S_A of the cylindrical antenna both $(\hat{n}\times E_1)$ and $(\hat{n}\times E_2)$ are constant around each cross section because of rotational symmetry. Furthermore, since $E_\phi=0$,

$$\hat{n}\times E=-\hat{\theta}\times E=-(\hat{\theta}\times\hat{r})E_r -(\hat{\theta}\times\hat{\theta})E_\theta=\hat{\phi}E_r. \tag{B8}$$

Hence, since $H=\hat{\phi}H_\phi$ for both fields,

$$\int_{SA}\{H_2\cdot(\hat{n}\times E_1)-H_1\cdot(\hat{n}\times E_2)\}d\sigma = \int_{r=a}^{r=a+h}E_{1r}dr\int_{\phi=0}^{\phi=2\pi}H_{2\phi}R_0d\phi - \int_{r=a}^{r=a+h}E_{2r}dr\int_{\phi=0}^{\phi=2\pi}H_{1\phi}R_0d\phi, \tag{B9}$$

where R_0 is the small radius of the antenna. However, from (B4) the contour integral of the

magnetic vector H_1 around the periphery of the antenna equals the total current in the antenna. Accordingly,

$$\int_{r=a}^{a+h} E_{2r}dr \int_{\phi=0}^{2\pi} H_{1\phi}R_0 d\phi = \int_{r=a}^{a+h} E_{2r}I_{1A}(r)dr, \quad \text{(B10)}$$

where $I_{1A}(r)$ is the total current in the antenna. Since the surface S_A (which is the surface of the antenna for E_1H_1) encloses only empty space for E_2H_2, the same contour integral with H_2 instead of H_1 equals only the time rate of change of $\epsilon_0 E_{2r}$ integrated across the small cross-sectional area πR_0^2. Since R_0 can be made as small as required, this quantity can be made negligible so that

$$\int_{r=a}^{r=a+h} E_{1r}dr \int_{\phi=0}^{\phi=2\pi} H_{2\phi}R_0 dr = 0. \quad \text{(B11)}$$

Substitution of (B7), (B10) and (B11) in (B6) gives

$$\int_0^{\pi} E_{2\theta}I_{1S}(\theta)ad\theta = \int_{r=a}^{r=a+h} E_{2r}I_{1A}(r)dr. \quad \text{(B12)}$$

This is the desired integral equation for determining $I_{1S}(\theta)$. $I_{1A}(r)$ is assumed given, $E_{2\theta}$ is the prescribed field given in (B5), and E_{2r} is determined from the vector $E_{2\theta}$ which satisfies the field equations and the boundary conditions (B5).

APPENDIX C

Linear Antenna and Sphere

The solution of the integral equation (B12) for the current $I_{1S}(\theta)$ on a metal sphere on which is erected a linear radiator with current $I_{1A}(r)$ is carried out as follows. Due to the orthogonality of the Legendre polynomials, that is,

$$\int_{-1}^{1} P_m(\mu)P_n(\mu)d\mu = [2/(2n+1)]\delta_{mn}, \quad \text{(C1)}$$

where $\delta_{mn} = \begin{cases} 1 \text{ if } m=n \\ 0 \text{ if } m\neq n \end{cases}$ is the Kronecker delta, it is possible to expand a function in terms of Legendre polynomials. Without limiting the generality the current $I_{1S}(\theta)$ on the surface of the sphere can be represented as follows

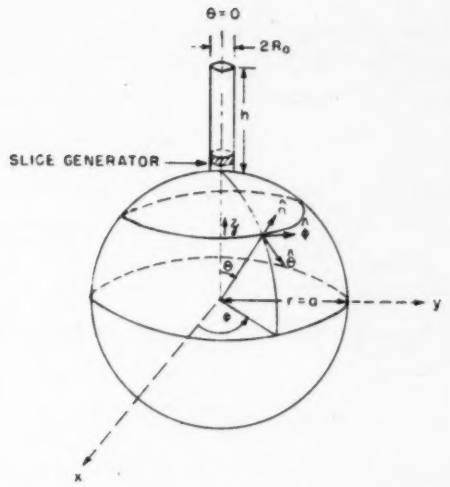


FIG. 11. The geometry of sphere excited by dipole.

$$I_{1S}(\theta) = \sum_{n=0}^{\infty} B_n P_n(\cos\theta). \quad \text{(C2)}$$

B_n are independent of θ and later it will be shown that they are functions of the radius a of the sphere and the height h of the antenna. Physically (C1) means that the total current crossing the parallel of latitude θ is the sum of all the so-called modal currents crossing this parallel of latitude; n is the order of the mode. As n increases, the amplitude of the current in this mode decreases so that it is possible to obtain a rather precise evaluation of $I_{1S}(\theta)$ by summing over the first twelve modes. Mathematically this means that (C1) converges for $0 \leqq \theta \leqq \pi$ rapidly enough so that terms of higher order than twelve contribute very little to the sum.

Expressed in appropriate coordinate form (B12) becomes the following for $r=a$:

$$\int_a^{a+h} E_{2r}I_{1A}(r)dr = \int_0^{\pi} E_{2\theta} \sum_{n=0}^{\infty} B_n P_n(\cos\theta) \cdot ad\theta. \quad \text{(C3)}$$

Furthermore, as was previously stated, E_{2r}, $E_{2\theta}$, and $H_{2\phi}$ are the only non-zero components of the field due to currents on the sphere; this means that the field is transverse magnetic TM.

It has been shown[9] that a TM field is uniquely derivable from a function u. Moreover, u satisfies the scalar wave equation:

$$\nabla^2(u/r)+k^2(u/r)=0, \tag{C4}$$

where $k=\omega/c$. The general solution of (C4) is

$$u=\sum_{m=0}^{\infty} A_m P_m(\cos\theta)\rho_m(kr), \tag{C5}$$

where $P_m(\cos\theta)$ is the Legendre polynomial of order m, and

$$\rho_n(x)=(\pi x/2)^{\frac{1}{2}}H_{m+\frac{1}{2}}^{(2)}(x)$$

is the weighted spherical Hankel function of the second kind. The components of the field are obtained from

$$\left.\begin{aligned} E_{2r}&=[k^2+(\partial^2/\partial r^2)]u,\\ E_{2\theta}&=(1/r)(\partial^2/\partial r\partial\theta)u,\\ H_{2\phi}&=-i(k/r)(\partial u/\partial\theta).\end{aligned}\right\} \tag{C6}$$

From (C5) and (C6), $E_{2\theta}$ is evaluated as follows:

$$E_{2\theta}=-(k/a)\sum_{m=0}^{\infty} A_m P_m^{1}(\cos\theta)\rho_m'(ka), \tag{C7}$$

where $P_m^{1}(\cos\theta)$ is the associated Legendre function. In order to evaluate the A_m in (C7), (B5) may be written in the form

$$\begin{aligned} E_{2\theta}&=U\sin\theta P_n(\cos\theta)\\ &=U[[P_{n+1}^{1}(\cos\theta)-P_{n-1}^{1}(\cos\theta)]/(2n+1)]. \end{aligned} \tag{C8}$$

Equating the coefficients of $P_m^{1}(\cos\theta)$ in (C7) and (C8) leads to

$$A_{n+1}=-aU/(2n+1)\cdot k\cdot\rho_{n+1}'(ka), \tag{C9}$$

$$A_{n-1}=+aU/(2n+1)\cdot k\cdot\rho_{n-1}'(ka). \tag{C10}$$

Substitution of these values of A_m in the first of Eqs. (C6) results in the following expression for E_{2r}:

$$E_{2r}=U\left[k^2+\frac{\partial^2}{\partial r^2}\right]\left[\frac{aP_{n-1}(\cos\theta)\rho_{n-1}(kr)}{(2n+1)\cdot k\cdot\rho_{n-1}'(ka)} -\frac{aP_{n+1}(\cos\theta)\rho_{n+1}(kr)}{(2n+1)\cdot k\cdot\rho_{n+1}'(ka)}\right]. \tag{C11}$$

[9] P. Frank and R. v. Mises, *Diff. Gleichungen d. Physik*, (Mary S. Rosenberg, New York) p. 872.

Now let it be assumed that the antenna current is

$$I_{1A}(r)=I_{\max}\sin k(d-r). \tag{C12}$$

With (C12) the left side of (C3) may be computed using (C11):

$$\begin{aligned} \int_a^{a+h} E_{2r}I_{1A}(r)dr&=\frac{UaI_{\max}}{(2n+1)\cdot k\cdot\rho_{n-1}'(ka)}\\ &\times\int_a^{a+h}\sin k(d-r)\cdot\left(k^2+\frac{\partial^2}{\partial r^2}\right)\rho_{n+1}(kr)dr\\ &-\frac{UaI_{\max}}{(2n+1)\cdot k\cdot\rho_{n+1}'(ka)}\int_a^{a+h}\sin k(d-r)\\ &\cdot\left(k^2+\frac{\partial^2}{\partial r^2}\right)\rho_{n+1}(kr)dr. \end{aligned}$$

This integration is carried out along the antenna, i.e., $\theta=0$, and since $P_n(1)=1$, the Legendre polynomials appearing in (C11) equal unity. Explicitly, the integration yields

$$\begin{aligned} \int_a^{a+h} E_{2r}I_{1A}(r)dr&=\frac{UaI_{\max}}{2n+1}\\ &\times\left[\frac{\rho_{n-1}(kd)}{\rho_{n-1}'(ka)}-\cos kh\frac{\rho_{n-1}(ka)}{\rho_{n-1}'(ka)}\right.\\ &\left.-\frac{\rho_{n+1}(kd)}{\rho_{n+1}'(ka)}+\cos kh\frac{\rho_{n+1}(ka)}{\rho_{n+1}'(ka)}\right]. \end{aligned} \tag{C13}$$

With (C8) the right-hand side of (C3) becomes:

$$\begin{aligned} \int_0^{\pi} E_{2\theta}\sum_{n=0}^{\infty} B_n P_n(\cos\theta)\cdot ad\theta&\\ =B_n\cdot[U2a/(2n+1)]&. \end{aligned} \tag{C14}$$

Equating (C13) and (C14), the following expression for B_n is obtained:

$$\begin{aligned} B_n=\frac{I_{\max}}{2}&\left[\frac{\rho_{n-1}(kd)}{\rho_{n-1}'(ka)}-\cos kh\frac{\rho_{n-1}(ka)}{\rho_{n-1}'(ka)}\right.\\ &\left.-\frac{\rho_{n+1}(kd)}{\rho_{n+1}'(ka)}+\cos kh\frac{\rho_{n+1}(ka)}{\rho_{n+1}'(ka)}\right]. \end{aligned} \tag{C15}$$

By substituting B_n from (C15) into (C2), $I_{1S}(\theta)$ is determined.

Spurious Signals Caused by Noise in Triggered Circuits*

DAVID MIDDLETON
Cruft Laboratory, Harvard University, Cambridge, Massachusetts
(Received January 23, 1948)

The expected number of peaks/sec. (n_{s+N}), caused by noise or by a signal and noise, in excess of a given amplitude, or triggering level, is determined for broad-band noise (a) and for the envelope of a narrow-band disturbance (b). The signals considered here are pulses of various shapes (rectangular, triangular, trapezoidal), or a carrier modulated by these pulses. The expected number of peaks per second is directly proportional to the band width of the noise, and depends in some degree (20–30 percent) on filter shape, the gaussian response yielding a larger value of n_{s+N} than the rectangular, for the same total input power. Further, n_{s+N} is greatest when the triggering level is the most probable amplitude. Pulse shape is not significant when the pulses are on only a small fraction of the total period of observation, but becomes so if (1) the time of observation is comparable to the pulse length, and (2) the pulses have sharp leading edges. The latter gives rise to a transient effect which is not ignorable. This transient contribution does not appear in the envelope of narrow-band waves, as long as the maximum time-rate of change of the wave form of the envelope is much less than the central or resonant frequency of the *IF*, i.e., as long as the concept of "envelope" is meaningful. Rectification does not change n_{s+N}, provided that there is no appreciable distortion of the envelope by video- or audio-filters and that the triggering level of the input is passed by the detector.

Improvement, defined in ways appropriate to the systems in question, of one circuit which can tolerate A_1 spurious signals/sec. in excess of a given triggering level over another circuit, which can tolerate only A_2 such signals ($A_1 > A_2$), is calculated for broad-band signal and noise, and for narrow-band waves following half-wave linear or quadratic rectification. In general, the improvement is slight (a few db) if the trigger levels of the two circuits are not close to the most probable value of the amplitude and if the ratio of the number of spurious signals allowed/sec. to the band width of the disturbance is small. Significant improvement follows only when the maximum tolerable number of triggering peaks/sec. is the same order of magnitude as the band width of the noise. Tables and figures illustrating the above are included in Part I; the mathematical derivations are contained in Part II.

I. DISCUSSION

1. Introduction

IN all instances in which signals of a certain magnitude are used to trigger a circuit, noise accompanies the triggering signal and may give rise to spurious responses as a result of the presence of noise peaks in excess of the excitation level, particularly if the signal voltage is comparable to the r.m.s. noise amplitude. Thus it is important to determine how many noise peaks per second, or how many on the average in a given time interval, will exceed the threshold level. Further, we may ask to know the improvement in using a circuit which can tolerate A_1 spurious pulses per second over another circuit which can tolerate only A_2 ($<A_1$) spurious pulses per second, the various filter widths and detector characteristics, etc. being essentially the same. In other words: how much greater is the sensitivity of circuit No. 1 over No. 2, given A_1 and A_2? This type of consideration and question clearly applies to all systems that depend for their performance on the knowledge of the number or average number of pulses received in a given time interval. Of course, the signals need not be in the form of pulses; for example, they may be sinusoidal, but the former are most widely used for trigger operations.

The analytical treatment of our problem is based on the observation that the expected number of peaks in excess of a certain value V_0 in the time interval (t_0, t_0+t) is the same as the expected number of crossings of the level V_0 with positive (*or* with negative) slope in the range (t_0, t_0+t), since once this level is exceeded, we have a triggering pulse. It is also clear that this number of crossings is not the same as the number of maxima, since the minimum between two successive peaks may lie above the level V_0. However, when V_0 is large, it may be shown[1]

* The research reported in this document was made possible through support extended Cruft Laboratory, Harvard University, jointly by the Navy Department (Office of Naval Research) and the Signal Corps, U. S. Army, under Contract N5ori-76, T.O. I.

[1] S. O. Rice, "Mathematical analysis of random noise," Bell Syst. Tech. J. **24**, 46 (1945). See Section 3.6, Eq. (3.6–11) in particular.

that we do approach the condition that the number of peaks or crossings is, on the average, equal to the number of maxima.

Rice[2] and Kac[3] have obtained, independently, expressions for the expected number of zeros ($V_0=0$) of a random wave. Our results (see Part II) are a generalization to the cases where $V_0 \neq 0$ and where there is a signal along with the random disturbance (Sec. 1, Part II). The expected number of peaks in excess of V_0, in the interval t_0, t_0+t follows from (II-1.5), and is

$$N_{\sigma+N}(V_0, t) = \int_{t_0}^{t_0+t} dt \int_0^\infty \dot{V} W_1(V_0, \dot{V}; t) d\dot{V}$$

$$= \int_{t_0}^{t_0+t} p(V_0; t) dt, \quad \text{(I-1.1)}$$

while the expected number of peaks per second is

$$n_{\sigma+N}(V_0, t) = N_{\sigma+N}(V_0, t)/t$$

$$= \frac{1}{t} \int_{t_0}^{t_0+t} p(V_0; t) dt. \quad \text{(I-1.2)}$$

The subscript σ refers to a possible signal and N to the noise. Here $W_1(V_0, \dot{V}; t)$ is the probability density for the variables $V(=V_0)$ and $\dot{V}$, and $p(V_0; t)$ is the probability density for the number of peaks exceeding V_0, both of which have been evaluated in Part II for a number of important cases. When the random process is stationary, i.e., the mechanism underlying the process does not change with time, we see at once that

$$N_{\sigma+N}(V_0, t) = t \cdot p(V_0)$$

and

$$n_{\sigma+N}(V_0) = p(V_0), \quad \text{(I-1.3)}$$

from (I-1.1) and (I-1.2).

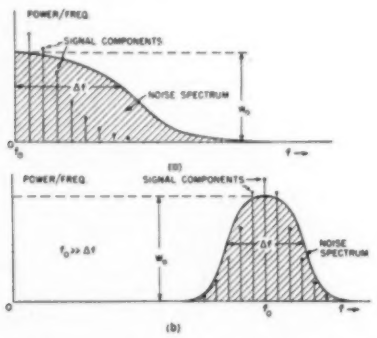


FIG. 1. Typical spectra of broad-band (a) and narrow-band (b) signal and noise waves.

Before turning to a consideration of specific results we note that there are two types of noise or noise and signal that are of interest here. These are (a) broad-band noise, the width Δf of whose spectrum between half-power points is equal to or greater than the resonant frequency, and (b) narrow-band noise, whose spectral width Δf is much less than the central or resonant frequency f_0.[4] In either case, an accompanying signal is assumed to be spectrally of the same order of magnitude as the noise. These remarks are illustrated in Fig. 1a, b; here W_0 denotes a spectral intensity (units: power/frequency). Type (a) spectra occur more frequently in acoustical problems, while type (b) is of greater interest in the reception of radio or radar signals.

When the disturbance consists of narrow-band noise and signal, for example a pulse-modulated carrier, it is usually rectified first, and the low frequency part[5] is then applied to trigger the circuit in question. Rectification does not destroy the validity of our results as long as the low-frequency filter following the detector is sufficiently wide to prevent any appreciable distortion of the detected envelope because of frequency selection, and provided, of course, that the rectifier passes voltages in excess of the triggering level. This is not surprising, since detection under these conditions merely distorts the wave without changing the *number* of peaks greater than the trigger level; relative wave shape is preserved, although the detector characteristic may be quite non-linear. Equations (I-1.1)–(I-1.3) may still be used, except that V_0 is now replaced by R_0, the triggering level in terms of the *undetected* envelope of the incoming signal and noise, and $\dot{V}$ by $\dot{R}$, the time rate of change of this envelope. The formal justification of these modifications

[2] S. O. Rice, Am. J. Math. **61**, 409 (1939), and see also Eq. (3.3–5), reference 1.

[3] M. Kac, Bull. Am. Math. Soc. **49**, 314 (1943). See reference 8.

[4] We consider here symmetrical spectra only.

[5] See D. Middleton, "Some general results in the theory of noise through nonlinear devices," Quart. App. Math. **5**, 445, Section 3, 1948, for the case of a signal as well as noise.

TABLE I.

Type of noise spectrum	Analytic form	b_0	b_2	$\rho=(b_2/b_0)^{\frac{1}{2}}$	Δf vs ω_F
1. *Rectangular*					
(a) broad band:	$w(f)=\begin{matrix}0, f>f_{F1}\\ W_0,\ 0<f<f_{F1}\end{matrix}$	$\frac{W_0\omega_{F1}}{2\pi}$	$\frac{W_0\omega^3_{F1}}{6\pi}$	$\omega_{F1}/(3)^{\frac{1}{2}}$	$\Delta f_1=\frac{\omega_{F1}}{2\pi}$
(b) narrow band:	$w(f)=\begin{matrix}0, f>f_0+f_{F2};\ f<f_0-f_{F2}\\ W_0,\ f_0-f_{F2}<f<f_0+f_{F2}\end{matrix}$	$\frac{W_0\omega_{F2}}{\pi}$	$\frac{W_0\omega^3_{F2}}{3\pi}$	$\omega_{F2}/(3)^{\frac{1}{2}}$	$\Delta f_2=\frac{\omega_{F2}}{\pi}$
2. *Gaussian*					
(a) broad band:	$w(f)=W_0e^{-\omega^2/\omega^2_{F1}}$	$W_0\omega_{F1}/4(\pi)^{\frac{1}{2}}$	$\frac{W_0\omega^3_{F1}}{8(\pi)^{\frac{1}{2}}}$	$\omega_{F1}/(2)^{\frac{1}{2}}$	$\Delta f_1=\frac{\omega_{F1}}{4\pi}$
(b) narrow band:	$w(f)=W_0e^{-(\omega-\omega_0)^2/\omega^2_{F2}}$	$W_0\omega_{F2}/2(\pi)^{\frac{1}{2}}$	$\frac{W_0\omega^3_{F2}}{4(\pi)^{\frac{1}{2}}}$	$\omega_{F2}/(2)^{\frac{1}{2}}$	$\Delta f_2=\frac{\omega_{F2}}{2\pi}$
3. *Optical*					
(a) broad band:	$w(f)=W_0/(1+\omega^2/\omega^2_{F1})$	$\frac{W_0\omega_{F1}}{4}$	∞	∞	$\Delta f_1=\frac{\omega_{F1}}{4}$
(b) narrow band:	$w(f)=W_0\Big/\left(1+\left(\frac{\omega-\omega_0}{\omega_{F2}}\right)^2\right)$	$\frac{W_0\omega_{F1}}{2}$	∞	∞	$\Delta f_2=\frac{\omega_{F2}}{2}$

is given in Part II, Section 2, and consequently (I-1.1)–(I-1.3) remain unchanged.[6]

2. Results and Conclusions

Let us consider the broad-band case (a) first, and let us restrict the discussion to include, at most, rectangular pulses of amplitude $V_\sigma=A_0$ and sufficiently long so that the transient effects of the vertical leading edge can be ignored. Expressions for a general signal are given in Part II, cf. (II-3.5). For the intervals that the signal is on or off the process is stationary, and from (I-1.3) and (II-3.14) we have at once for the expected number of peaks per second in excess of V_0:

$$n_{(\sigma)+N}(V_0)=\rho\left(\frac{1}{2\pi}e^{-v_0^2}\right)=p(V_0), \quad \text{(I-2.1)}$$

where ρ is the *band-width factor*, proportional to the width Δf of the noise spectrum and is given by

$$\rho^2\equiv\frac{b_2}{b_0}=\frac{\int_0^\infty(\omega-\omega_0)^2w(\omega)d\omega}{\int_0^\infty w(\omega)d\omega},$$
$$\omega=2\pi f, \quad \omega_0=2\pi f_0. \quad \text{(I-2.2)}$$

When the signal is on or off

$$v_0=(V_0-V_\sigma)/(2b_0)^{\frac{1}{2}}, \quad \text{or} \quad v_0=V_0/(2b_0)^{\frac{1}{2}}, \quad \text{(I-2.3)}$$

respectively. The presence of the d.c. V_σ simply serves to "boost" the noise, so that the effective triggering level is now V_0-V_σ. The quantity b_0 is the mean noise power passed by the filter, and $w(f)$ is the mean noise power spectrum; ω_0 is zero for the broad-band disturbances treated here.

The general result for the expected number of peaks in excess of R_0 in the case of the envelope of narrow-band noise and/or a signal has been obtained in Part II; see Sections 4 or 5, in particular Eqs. (II-4.9) and/or (Eq. II-5.11). Where there is no signal we have

$$n_N(R_0)=\rho\left(\frac{r_0}{(\pi)^{\frac{1}{2}}}e^{-r_0^2}\right), \quad r_0\equiv R_0/(2b_0)^{\frac{1}{2}}, \quad \text{(I-2.4)}$$

and for an unmodulated carrier

$$n_{\sigma+N}(R_0)=\rho\left[\frac{r_0}{(\pi)^{\frac{1}{2}}}e^{-r_0^2-a_0^2}I_0(2a_0r_0)\right],$$
$$a_0\equiv A_0/(2b_0)^{\frac{1}{2}}, \quad \text{(I-2.5)}$$

[6] Care must be exercised in dealing with signals that have a large time rate of change, of which the idealized, rectangular pulse used here is an extreme example. Signals with steep leading edges contain transients whose contributions to the expected number of peaks exceeding V_0 often cannot be ignored. When we deal only with the envelope these transients do not appear, to a first approximation, since they consist chiefly of frequencies outside the spectral limits of the envelope. A rigorous treatment would include the effects of finite filter widths, but this is a refinement unnecessary for the purposes of the present work, since pulse shape in most instances does not have a critical effect on the number of peaks exceeding R_0.

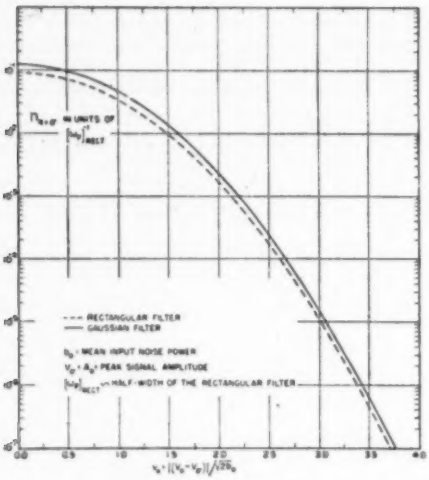


FIG. 2. Expected number of noise peaks per second in excess of V_0 for broad-band noise.

where A_0 is the peak amplitude of the carrier. Equation (I-2.5) applies equally well for a carrier modulated by a rectangular pulse[6] for the interval the pulse is on, while (I-2.4) holds during the time the pulse is off.

Although the rectangular pulse is an idealization, other pulse shapes, such as the gaussian, triangular, trapezoidal, etc., give values of $n_{\sigma+N}$ that are not much different from those obtained in the rectangular case. Only for strong signals (relative to the noise) is the number of peaks in the rectangular case greater than for the other pulses (of the same maximum amplitude and width between half-power points), since in the former the noise is "boosted" during the entire duration of the pulse, while for the others this is not true. As mentioned in Section I, pulse shape is not critical in its effect on $n_{\sigma+N}$, and is of negligible importance when the duty-cycle τ is a small fraction of the repetition period T, the most common case in practice. Thus, the assumption of a rectangular pulse, while not strictly valid, is quite good enough for most uses.

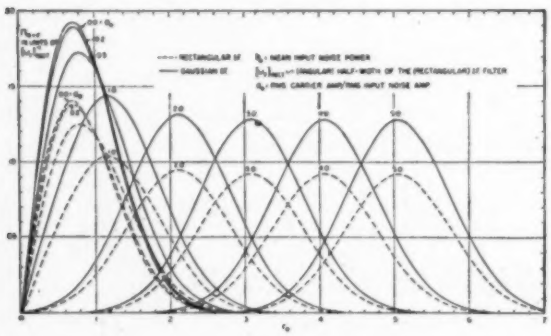


FIG. 3. The expected number of noise peaks per second in excess of r_0 (≡triggering level/rms noise level). The d.c. component is included.

The spectral width of the noise significantly affects the number of peaks exceeding our trigger level V_0 or R_0. In fact, for noise alone and for the types of signal mentioned above, *$n_{(\sigma)+N}$ is directly proportional to the band-width Δf of the noise mean power spectrum*, whether broad- or narrow-band. From the physical point of view we would expect this, since the number of crossings at any level depends not only on the probability that the wave will reach or exceed such a value but also on how often the wave fluctuates, which, on the average, depends on the number of times the disturbance passes through zero. This number, in turn, is of the same order of magnitude as the mean frequency of the noise and hence is proportional to its band width. The constant of proportionality depends on spectral shape, although the variation from one distribution to another is not very pronounced. Table I lists data on several types of spectra considered in this paper.

For purposes of comparison we normalize our spectra so that they all have the same maximum intensity W_0 and equal power; ω_F is accordingly different in the various cases. From the table we see immediately that

$$(\omega_F)_{\text{gauss}} = \frac{2}{(\pi)^{\frac{1}{2}}}(\omega_F)_{\text{rect.}} = (\pi)^{\frac{1}{2}}(\omega_F)_{\text{opt.}},$$

where

$$(\omega_F)_{(a)} = 2(\omega_F)_{(b)}. \qquad \text{(I-2.6)}$$

Then the band-width factors ρ are related by

$$(\rho)_{\text{gauss}} = \left(\frac{2}{\pi}\right)^{\frac{1}{2}}(\omega_F)_{\text{rect.}} = \left(\frac{6}{\pi}\right)^{\frac{1}{2}}\rho_{\text{rect.}};$$

$$\rho_{\text{rect.}} = \omega_F/(3)^{\frac{1}{2}}. \qquad \text{(I-2.7)}$$

It is interesting to note that for the "optical" spectrum (single-tuned circuits), (3) Table I, ρ becomes infinite. The spectrum does not tail off to zero intensity with sufficient rapidity. In actual practice, however, this catastrophe does not occur, as damping or parasitic effects insure convergence; the ideal "optical" spectrum is

modified so that the contributions to the bandwidth factor ρ for frequencies well away from resonance can be completely ignored (see reference 1, Section 3.3 for further details).

Curves showing the variation of the expected number of peaks per second with trigger level are included in Figs. 2 and 3. In general, this number is greatest at triggering levels for which the probability density of the amplitude or envelope is a maximum, or, approximately, for levels which are equal to the mean value of the disturbance. In the case of broad-band noise these points are $V_0=0$, or $V_0=V_\sigma$ if there is a steady component (Fig. 2). For the envelope R of narrow-band waves the mean value ($\bar{R}$) is always greater than zero, as Fig. 3 indicates. Note, moreover, that for triggering levels R_0 below $\bar{R}$ there are actually fewer peaks/sec. than at $R_0=\bar{R}$; maxima and minima above this level, and in the time interval between an upward crossing of R_0 and the next downward one, merely act as a single peak exceeding R_0 as far as triggering effects are concerned. Gaussian noise spectra, with the same power and maximum intensity as rectangular ones, always give a greater number of noise peaks.

Improvement

We are now able to determine the improvement (db) in the minimum allowable trigger level of one circuit which can tolerate A_1 spurious pulses per second over a simpler one which can tolerate only A_2 spurious pulses ($A_2 \leq A_1$). Our circuits are assumed to have otherwise identical characteristics: $(b_0)_1=(b_0)_2$ and $\rho_1=\rho_2$.

There are various ways in which "improvement" may be defined, depending on the mode of operation of our circuits. For example, we may use the following definitions:

Let

$I_1=$ {output level (after detection, if any) for which the expected number of *spurious* pulses/sec. for circuit No. 1 is A_1}.

$I_2=$ {output level, as above, for which the expected number of spurious pulses/sec. is A_2 ($A_2 \leq A_1$)}.

Then the improvement is

$$\Gamma \equiv 20 \log_{10}(I_2/I_1) \text{ db}; \quad (I_2 \geq I_1;\ A_2 \leq A_1). \quad \text{(I-2.8)}$$

If the d.c. is excluded, as is frequently the case in audio or video filtering following a detector, our definition (I-2.8) is modified to

$$\Gamma(-\text{d.c.}) \equiv 20 \log_{10}\left[\frac{I_2-(I_{\text{d.c.}})_2}{I_1-(I_{\text{d.c.}})_1}\right] db (I_2 \geq I_1). \quad \text{(I-2.9)}$$

The relation between the expected number of spurious pulses/sec. (A_1, A_2) and the triggering levels (I_1, I_2), (R_{01}, R_{02}), etc., in our definitions of improvement may be obtained from (I-1.1), (I-1.2), and (II-3.16). For pulses (of any shape) whose duration τ is short compared to the repetition period T ($T \gg \tau$), so that the effect of the signal is negligible, we have from (I-2.8) and (II-3.19) on solving for V_{01} and V_{02},

$$\Gamma = 20 \log_{10}\left(\frac{V_{02}}{V_{01}}\right) = 10 \log_{10}\left(\frac{\log_{10} 2\pi y_2}{\log_{10} 2\pi y_1}\right) db \quad \text{(I-2.10)}$$

for unrectified, broad-band waves. Here

$$y \equiv [A+\alpha(v_0, a_0; T)]/\rho \quad (\leq 1/2\pi), \quad \text{(I-2.11)}$$

where α is a correction term caused by the transient effects of the leading edge of the pulse (see II-3.18). Figure 4 shows Γ as a function of y_2.

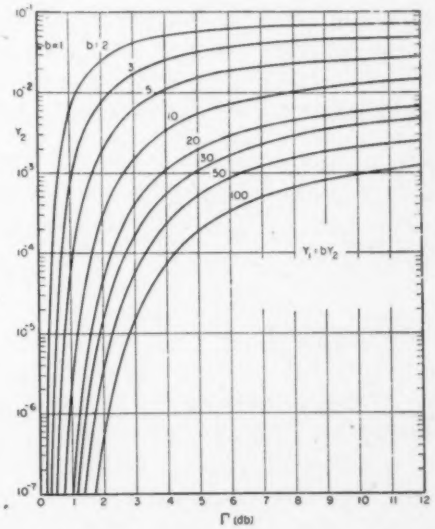


FIG. 4. Here y_1 and y_2 are factors depending on the expected number of spurious pulses per second, on the period of the modulation, on the input carrier strength, noise level, etc. The precise relationship is given in Eq. (I-2.11) and the "improvement" (db) of circuit No. 1 over No. 2 for broad-band noise and short pulses follows from Eq. (I-2.10).

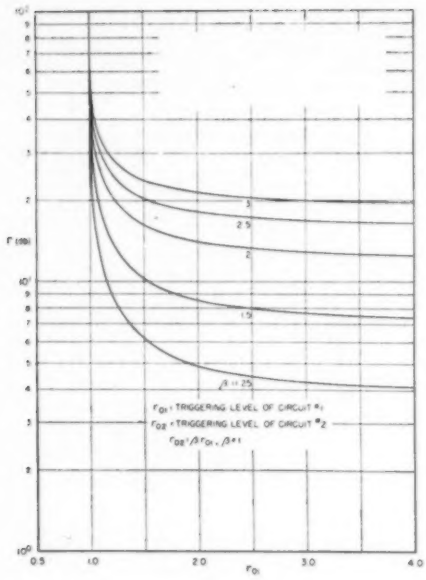


FIG. 5. Improvement (db) of circuit No. 1 over No. 2 after half-wave linear rectification of narrow-band noise and short pulses d.c. component of the envelope removed.

A more important case is that of narrow-band signal and noise triggering a circuit after rectification by a half-wave linear or quadratic detector. For the moment we assume that the d.c is passed by the video or audio filter. The total low frequency and d.c. outputs of the respective rectifiers are[7]

$$\text{(linear)}: \; I_{ll}=\beta R/\pi; \; I_{\text{d-c}}=\beta(b_0/2\pi)^{\frac{1}{2}}, \quad \text{(I-2.12a)}$$

$$\text{(quadratic)}: \; I_{ll}=\beta R^2/4; \; I_{\text{d-c}}=\beta b_0/2. \quad \text{(I-2.12b)}$$

Here β is a tube constant. Again the signal is assumed to be on a negligible portion of the period. Our result (I-2.4) may thus be used in (I-2.8) to give us

$$\Gamma_{(\text{lin.})}=20\log_{10}\left(\frac{r_{02}}{r_{01}}\right)$$

$$=20\log_{10}\left(\frac{y_2\exp(r_{02}^2)}{y_1\exp(r_{01}^2)}\right)=\tfrac{1}{2}\Gamma_{\text{quad.}}(db),$$

$$(r_{02}\geq r_{01};\; y_2\leq y_1), \quad \text{(I-2.13)}$$

where now the transient term $\alpha(v_0, a_0; T)$ in y (cf. I-2.11) has the value $1/T$. When the d.c. component is filtered out, which is the more common case in practice, we must use the definition (I-2.9), which becomes with aid of (I-2.12)

$$\Gamma_{(\text{lin.})}(-\text{d.c.})=20\log_{10}\left(\frac{r_{02}-(\pi)^{\frac{1}{2}}/2}{r_{01}-(\pi)^{\frac{1}{2}}/2}\right)$$

$$=20\log_{10}\left(\frac{y_2\exp(r_{02}^2)-1/2}{y_1\exp(r_{01}^2)-1/2}\right), \quad \text{(I-2.14a)}$$

and

$$\Gamma_{(\text{quad.})}(-\text{d.c.})=20\log_{10}\left(\frac{r_{02}^2-1}{r_{01}^2-1}\right)$$

$$=20\log_{10}\left(\frac{y_2^2\exp(2r_{02}^2)-1/\pi}{y_1^2\exp(2r_{01}^2)-1/\pi}\right). \quad \text{(I-2.14b)}$$

The various values of r_0 are obtained most easily from a plot of $y=(A+\alpha)/\rho$ as a function of r_0, *viz*:

$$y=\frac{r_0}{(\pi)^{\frac{1}{2}}}\exp(-r_0^2).$$

From this relation (Fig. 7) we note that $y_{1,2}\leq(2\pi)^{-\frac{1}{2}}$, so that Γ must exceed zero.

Curves for the improvement following linear or quadratic detection are shown in Figs. 5 and 6 as a function of $r_{01}=r_{02}/\beta$, where β (≥ 1) relates

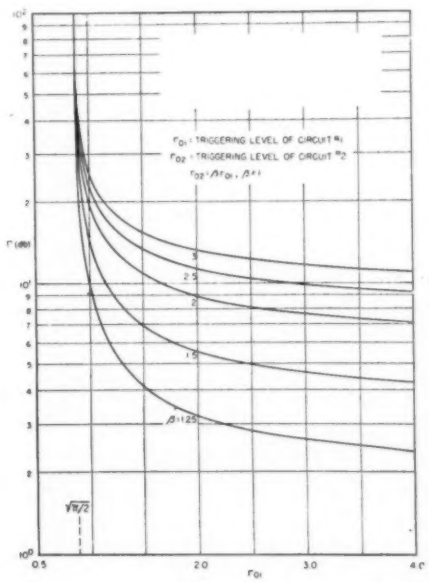


FIG. 6. Improvement (db) of circuit No. 1 over No. 2 after half-wave quadratic detection of narrow-band noise and short pulses, d.c. component of the envelope removed.

[7] Reference 1, Section 4.1. I_{ll} may also be obtained from (II-2.1).

the critical envelope levels r_{01} and r_{02} at which triggering takes place in the two circuits. We observe that, for the cases when the d.c. is passed, if we set our triggering level below the point of most probable amplitude ($\sim \bar{r}_0$), i.e., $r_{01} < (r_0)_{\max} < r_{02}$, the improvement *decreases* from its maximum value at $r_{01} = (r_0)_{\max}$ until r_{01} reaches a value for which $y_1 = y_2$, at which $\Gamma = 1$; r_{01} below this level gives a response for circuit No. 1 less favorable than for circuit No. 2. When the d.c. ($\bar{r}_0$) is filtered out, however, this situation cannot occur, since $(r_0)_{\max} \leq \bar{r}_0$, and $r_{01} \geq \bar{r}_0$.

In general, the improvement is not great (a few db) if the trigger levels of the two circuits are not close to the most probable value of the amplitude and if the ratio of the number of spurious signals allowed/sec. (A_1, A_2) to the bandwidth of the disturbance is small. This may be observed from Fig. 4, or from Figs. 5 and 6 used in conjunction with Fig. 7. Significant improvement (>10 db) follows only if circuit No. 1 can tolerate a much larger number of spurious pulses/sec. than circuit No. 2 and if this number is the same order of magnitude as the bandwidth factor ρ.

The author wishes to express his gratitude to Miss Marilyn Lang, who performed the computations.

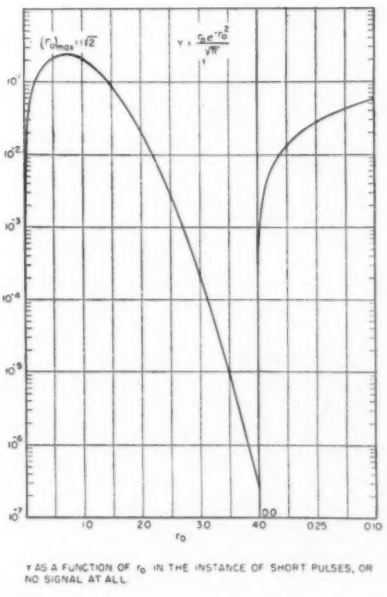


FIG. 7. The parameter y is given by Eq. (I-2.11) and has the same significance as in Fig. 4; r_0 is defined as in Fig. 3.

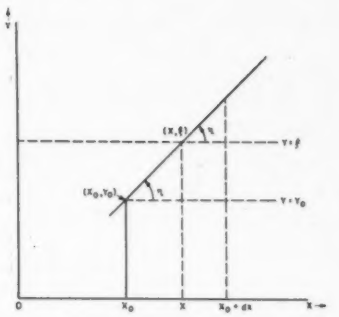


FIG. 8.

II. MATHEMATICAL SECTION

1. The Expected Number of Noise Peaks Exceeding a Given Magnitude

The expected number of noise peaks that exceed a given level may be determined in the following fashion. We begin by calculating the probability that a function

$$y = y(a_1 \cdots a_N ; b_1 \cdots b_N ; x) \qquad \text{(II-1.1)}$$

passes through the value ξ in the interval (x_0, x_0+dx) with positive (*or*, negative) slope. Here the a_i, b_i ($i = 1, 2, \cdots N$) are random parameters, possessing certain statistical properties (cf. Section 3). We choose dx so small that the parts of all but an ignorable (infinitesimal of higher order) fraction of the possible random curves lying in the strip (x_0, x_0+dx) may be regarded as straight lines. If $y = y_0$ at $x = x_0$ and passes through $y = \xi$ (at $x = x$) for $x_0 < x < x_0+dx$, the slope η of our line becomes

$$\eta = (\xi - y_0)/(x - x_0) ; \qquad \text{(II-1.2)}$$

see Fig. 8 above. Since $x_0 < x < x_0+dx$ we obtain the inequality

$$0 > (y_0 - \xi) > -\eta dx. \qquad \text{(II-1.3)}$$

Now let $W_1(y_0, \eta ; x_0)$ be the probability density function for the random variables y_0 and $\eta = \partial y/dx|_{x=x_0}$ for a specified value of $x = x_0$. Then the probability that y_0 and η satisfy our inequality (II-1.3), going through $y = \xi$ with positive slope in the interval (x_0, x_0+dx), is

$$dx \int_0^\infty \left[\int_{y_0=\xi-\eta dx}^{y_0=\xi} W_1(y_0, \eta ; x_0) dy_0 \right] d\eta$$

$$\equiv p(\xi ; x_0) dx, \qquad \text{(II-1.4)}$$

which gives us

$$dx\left[\frac{d}{d(\Delta x)}\int_0^\infty d\eta\int_{\xi-\eta\Delta x}^{\xi} W_1(y_0,\,\eta;\,x_0)dy_0\right]_{\Delta x\to 0}$$

$$=dx\left[\int_0^\infty d\eta\left\{-\frac{d}{d(\Delta x)}(-\eta\Delta x)\right\}\right.$$

$$\left.\times W_1(\xi-\eta\Delta x,\,\eta;\,x_0)\right]_{\Delta x\to 0}$$

after differentiation in the limits; and thus we find[8] **

$$p(\xi;\,x_0)=\int_0^\infty \eta W_1(\xi,\,\eta;\,x_0)d\eta. \quad \text{(II-1.5)}$$

In our present case we have $y=V(t)$, whose nature will be specified more precisely in the following sections. The expected number of crossings of $y=\xi=V_0$ by $V(t)$, with positive slope in the time interval $t-t_0$, or what is the same thing, the expected number of "peaks" in excess of V_0 in that interval, is given explicitly by (I-1.1) on a simple change of notation. The expected number of peaks per second is obtained from (I-1.2).

Our problem has now been reduced to calculating the probability density $p(V_0;t)$ for a random noise wave, or one that contains noise and a signal, and evaluating the integrals (I-1.1) or (I-1.2).

2. The Effect of Rectification

We examine first the case of broad-band noise; see Fig. 1a. Let $I(t)$ be the output (signal and noise) current after rectification by a device the Fourier transform of whose dynamic characteristic $g(V)$ is $f(iz)$. The output may be expressed as

$$I(t)=\frac{1}{2\pi}\int_{\mathbf{C}} f(iz)e^{izV(t)}dz, \quad \text{(II-2.1)}$$

where **C** is the familiar contour[8,9] extending along the real axis from $-\infty$ to $+\infty$ and is indented downward about a possible singularity at the origin. It is further understood that any video or audio filter following the detector is essentially uniform and infinitely wide for our purposes.

From (II-1.5) the probability density for the expected number of peaks in excess of an arbitrary output I_0 (≥ 0) is

$$p(I_0;\,t)=\int_0^\infty \dot{I}P(I_0,\,\dot{I};\,t)d\dot{I}, \quad \text{(II-2.2)}$$

where I and I_0 are related to V and V_0 by (II-2.1), and P is a density function analogous to W_1. Differentiating (II-2.1) we observe that since $V=V_0$, $dV=0$, we have finally

$$p(I_0;\,t)=\int_0^\infty \left(\frac{\partial I}{\partial V}\right)^2_{I=I_0}\dot{V}P(I_0,\,\dot{I};\,t)d\dot{V}. \quad \text{(II-2.3)}$$

Our assumption of no distortion because of frequency selection requires that the *total* probability for the number of peaks before and after detection be unity. This permits us to write

$$P(I_0,\,\dot{I},\,t)=W_1(V_0,\,\dot{V};\,t)\Big/\left.\frac{\partial(I,\,\dot{I})}{\partial(V,\,\dot{V})}\right|_{I=I_0,\,V=V_0}$$

$$=\left(\frac{\partial I}{\partial V}\right)^{-2}_{I=I_0}W_1(V_0,\,\dot{V};\,t). \quad \text{(II-2.4)}$$

When (II-2.4) is substituted into (II-2.3) the result is

$$p(I_0;\,t)=\int_0^\infty \dot{V}W_1(V_0,\,\dot{V};\,t)d\dot{V}$$

$$=p(V_0;\,t) \quad \text{(II-2.5)}$$

as expected.

For narrow-band noise (Fig. (1b)) and a signal in its immediate spectral vicinity the same general procedure may be used to show that for the low frequency output $I_{\text{L.F.}}$, i.e., the rectified envelope,

$$p([I_{\text{L.F.}}]_0;\,t)=\int_0^\infty \dot{R}W_1(R_0,\,\dot{R};\,t)d\dot{R}$$

$$=p(R_0;\,t), \quad \text{(II-2.6)}$$

[8] This is a generalization of Rice's result, Eq. (3.3-5), Bell Sys. Tech. J. **24**, 46 (1945), in which he computes the expected number of zeros, i.e., $\xi=0$. In this instance ($\xi=0$) results equivalent to those of Rice (see also Am. J. Math. **61**, 409 (1939)) have been obtained independently by M. Kac, Bull. Am. Math. Soc. **49**, 314 (1943) for somewhat less stringent conditions on the random parameters (a_i,b_i). See also M. Kac, Am. J. Math. **65**, 609 (1943).

** Since this was written, the result (II-1.5) has been stated by Rice, Bell Sys. Tech. J. **27**, 109 (1948), Eq. (2.1); see also the remarks in footnote 10 of that paper.

[9] See, for example, reference 1, Sec. 4.3, and Appendix 4A.

where R (≥ 0) is the input envelope, R_0 the level exceeded, and $W_1(R_0, \dot{R}; t)$ is the probability density obtained from $W_1(V_0, \dot{V}; t)$ on transforming from $(V, \dot{V})$-space to $(R, \dot{R})$-space.

3. Peak-Number Distribution Density for Signal and Broad-Band Noise.

Let our disturbance, consisting of a signal component $V_\sigma(t)$ and noise $V_N(t)$ be represented by

$$V(t) = V_N(t) + V_\sigma(t) \equiv Y_1 = X_1 + V_\sigma$$

$$= V_\sigma + \sum_{n=1}^{\infty} (a_n \cos\omega_n t + b_n \sin\omega_n t), \qquad \omega_n = 2\pi f_n, \quad \text{(II-3.1)}$$

and

$$\therefore \quad \dot{V}(t) = \dot{V}_N + \dot{V}_\sigma \equiv Y_2 = \dot{V}_\sigma + X_2$$

$$= \dot{V}_\sigma + \sum_{n=1}^{\infty} (-a_n\omega_n \sin\omega_n t + b_n\omega_n \cos\omega_n t), \quad \text{(II-3.2)}$$

in which a_n and b_n are the familiar random parameters[5] having the properties

$$\langle a_i b_j \rangle = 0; \quad \langle a_i \rangle = \langle b_j \rangle = 0;$$
$$\langle a_i a_j \rangle = \langle b_i b_j \rangle = \delta_j{}^i \Delta f w(f_i). \quad \text{(II-3.3)}$$

Here $\delta_j{}^i$ is Kronecker delta, such that $\delta_j{}^i = 0$, $i \neq j$; $\delta_i{}^j = 1$, $w(f_i)$ is the spectral intensity (mean power density) at the frequency f_i, and Δf is the separation between adjacent components in the Fourier development (see II-3.1, 3.2). The brackets $\langle\ \rangle$ indicate a statistical average. The quantities X_1 and X_2 are the gaussian random variables whose distribution $W_1(X_1, X_2)$ we seek. In terms of $Y_1 = V_0$, the level to be exceeded, and $Y_2 = X_2 + \dot{V}_\sigma$, we observe that

$$W_1(V_0, Y_2; t) = \frac{1}{2\pi(b_0 b_2)^{\frac{1}{2}}} \exp[-(V_0 - V_\sigma)^2/2b_0 - (Y_2 - \dot{V}_\sigma)^2/2b_2]; \quad \text{(II-3.4a)}$$

where

$$b_n \equiv \int_0^\infty \omega^n w(f) df. \quad \text{(II-3.4b)}$$

W_1 is now a function of t, in as much as $V_\sigma = V_\sigma(t)$ and $\dot{V}_\sigma = \dot{V}_\sigma(t)$ in the general case. Equation (II-1.5) or (II-2.5) then integrates to

$$p(V_0; t) = \frac{\rho}{2\pi} \exp[-(V_0 - V_\sigma)^2/2b_0] \times \{\exp(-\dot{V}_\sigma{}^2/2b_0\rho^2) + \dot{V}_\sigma(\pi/2b_0\rho^2)^{\frac{1}{2}} \times (1 + \Theta[\dot{V}_\sigma/(2b_0\rho^2)^{\frac{1}{2}}])\} \quad \text{(II-3.5a)}$$

or

$$= \frac{\rho}{(2\pi)^{\frac{1}{2}}} \exp[-(V_0 - V_\sigma)^2/2b_0] \times \left\{ \frac{\dot{V}_\sigma}{2(b_0\rho^2)^{\frac{1}{2}}} + \frac{1}{(2\pi)^{\frac{1}{2}}} \cdot \times {}_1F_1(-\tfrac{1}{2}; \tfrac{1}{2}; -\dot{V}_\sigma{}^2/2b_0\rho^2) \right\}, \quad \text{(II-3.5b)}$$

where $\rho \equiv (b_2/b_0)^{\frac{1}{2}}$. The quantity Θ is the error integral defined by

$$\Theta(x) = \frac{2}{(\pi)^{\frac{1}{2}}} \int_0^x \exp(-y^2) dy,$$

and tabulated, for example, in B. O. Pierce's, *A Short Table of Integrals* (Ginn and Company, New York, 1929). The expression ${}_1F_1$ is a confluent hypergeometric function, whose more important properties are listed in references 1 and 5.

When there is no signal, we obtain Kac's[3, 8] or Rice's result (reference 8, Eq. (3.3–13)), namely

$$p(V_0) = \frac{\rho}{2\pi} \exp(-V_0{}^2/2b_0) = \exp(-V_0{}^2/2b_0) \cdot \times [\tfrac{1}{2} \text{ the expected no. of zeros/sec.}]. \quad \text{(II-3.6)}$$

Since now the process is stationary, $p(V_0)$ is independent of t.

Notice in the general case (II-3.5), that the probability density becomes infinite ($\sim \dot{V}_\sigma$) for signals possessing an infinite time derivative. The physical significance of this point is discussed more fully as we consider (*a*) and (*b*) below.

a. D.C. Signal

Here the signal appears simply as a d.c. voltage. Then $\dot{V}_\sigma$ vanishes, *except* for the leading edge, marking the point (in time) at which the signal

is first applied. Let us assume for the moment that the change from no signal to a steady component of amplitude A_0 takes place linearly so that we have

$$\left.\begin{array}{lll} V_s(t)=A_0t/\tau_0; & 0\leq t\leq\tau_0; & \dot{V}_s=A_0/\tau_0 \\ =A_0 & t\geq\tau_0; & =0 \end{array}\right\}. \quad \text{(II-3.7)}$$

From (II-3.5) the probability density for the expected number of peaks/sec. in excess of V_0 in the transition region ($0\leq t\leq\tau_0$) is

$$p(V_0;t)=\frac{1}{2\pi\tau_0}\exp[-(v_0-a_0/\tau_0)^2] \times\{(\rho\tau_0)\exp[-a_0^2/(\rho\tau_0)^2] +a_0(\pi)^{\frac{1}{2}}[1+\Theta(a_0/(\rho\tau_0))]\}, \quad \text{(II-3.8a)}$$

where $v_0\equiv V_0/(2b_0)^{\frac{1}{2}}$ and $a_0\equiv A_0/(2b_0)^{\frac{1}{2}}$, and in the steady-state interval ($t\geq\tau_0$) we have

$$p(V_0)=(\rho/2\pi)\exp[-(v_0-a_0)^2]=(b_2/2\pi)^{\frac{1}{2}} \times(\text{probability density for noise distributed about the average value } A_0). \quad \text{(II-3.8b)}$$

The probability density is greater than it would be without signal, since the noise is "boosted" to a new average level (v_0-a_0).

Observe in (II-3.8a) that as $\tau_0\to 0$ ($t\to 0$ also, since $0\leq t\leq\tau_0$), $p(V_0;t)$ becomes infinite as τ_0^{-1}. The quantity $(\rho\tau_0)$, which is proportional to the ratio of filter width to signal width (freq.), however, remains finite and possibly different from zero as $\tau_0\to 0$. Usually $(\rho\tau_0)$ vanishes as $\tau_0\to 0$, since the amount of noise power is assumed to be constant, with non-zero intensity W_0; this means that $\rho=$a finite constant. It may be objected that to achieve the linear transient of our example (II-3.7) above, even when $\tau_0>0$, requires an infinite filter width. This is true, but the *slope* of the transient remains finite, and so $p(V_0;t)$ presents no singularities. In practice, moreover, our filters never pass the highest harmonics, with the result that the slope of the transient, or signal, is always bounded. The quantity $(\rho\tau_0)$, or an equivalent expression for waves of other shapes, never quite vanishes, although it may become very small in the case of filters with high resolution. In these instances the limiting form as $\tau_0\to 0$ is a good approximation, i.e., the ideal vertical edge is a safe model for the real, not-quite vertical one.

In an interval $T(>\tau_0)$ from the time the signal first goes on, the expected number of peaks per second exceeding V_0 is, from (I-1.2),

$$n_{s+N}(V_0)=\frac{\rho}{2\pi}\left[1-\frac{\tau_0}{T}\right]\exp[-(v_0-a_0)^2] +[4Ta_0(\pi)^{\frac{1}{2}}]^{-1}[\Theta(v_0)-\Theta(v_0-a_0)] \times(\rho\tau_0\exp(-a_0^2/(\rho\tau_0)^2) +a_0(\pi)^{\frac{1}{2}}[1+\Theta(a_0/\rho\tau_0)]). \quad \text{(II-3.9)}$$

For long times ($T\gg\tau_0$) the transient's effect becomes quite negligible, as we would expect, and $n_{s+N}(V_0)$ is now stationary and reduces to (II-3.8b). The precise condition when $p(V_0)$ may be used in place of $n_{s+N}(V_0)$ for signals of the above type may be obtained from (II-3.9) on requiring that $\rho/2\pi\exp[-(v_0-a_0)^2]$ exceed the remaining terms by a sufficient amount.

In the case of a vertical leading edge our result (II-3.9) assumes a simpler form since as $\tau_0\to 0$, $\rho\tau_0$ vanishes also, in consequence of the condition of finite noise power. We have accordingly

$$n_{s+N}(V_0)=\frac{\rho}{2\pi}\exp[-(v_0-a_0)^2] +\frac{1}{2T}[\Theta(v_0)-\Theta(v_0-a_0)]. \quad \text{(II-3.10)}$$

Again, for sufficiently long times the contribution of the second term proves negligible and $n_{s+N}(V_0)=p(V_0)$, Eq. (II-3.8b).

b. Signal As a Trapezoidal, Triangular, or Rectangular Pulse

Before considering a rectangular pulse let us evaluate $p(V_0;t)$ and n_{s+N} for a trapezoidal pulse of width τ, with the following analytic representation:

$$\left.\begin{array}{lll} V_s=A_0t/\tau_0, & 0\leq t\leq\tau_0, & \dot{V}_s=A_0/\tau_0 \\ =A_0, & \tau_0\leq t\leq\tau-\tau_0, & =0 \\ =A_0(\tau-t/\tau_0), & \tau-\tau_0\leq t\leq\tau, & =-A_0/\tau_0 \end{array}\right\}. \quad \text{(II-3.11)}$$

The rectangular pulse, of duration τ, may then be treated as the limiting form of the above as $\tau_0\to 0$. From (II-3.5) we see that the expressions for $p(V_0;t)$ are precisely those given in (II-3.8a)

and (II-8b); the former applies in the interval $(0 \leq t \leq \tau_0)$, and during the time $(\tau-\tau_0 \leq t \leq \tau)$ provided we replace t by $\tau-t$; the latter holds for the period $(\tau_0 \leq t \leq \tau-\tau_0)$. The expected number of peaks/sec. exceeding V_0 *while the pulse is on* follows at once from (I-1.2) and is

$$n_{s+N}(V_0)=\frac{\rho}{2\pi}\left(1-\frac{2\tau_0}{\tau}\right)\exp[-(v_0-a_0)^2]$$
$$+\frac{1}{2\tau a_0(\pi)^{\frac{1}{2}}}[\Theta(v_0)-\Theta(v_0-a_0)]$$
$$\times[(\rho\tau_0)\exp(-a_0^2/(\rho\tau_0)^2)$$
$$+a_0(\pi)^{\frac{1}{2}}\Theta(a_0/\rho\tau_0)]. \quad \text{(II-3.12)}$$

When the transients fore and aft are short compared to the total duration of the pulse $(\tau \gg 2\tau_0)$ we obtain the familiar result (II-3.8b).

An expression similar to (II-3.12) for triangular pulses may be deduced immediately on setting $\tau=2\tau_0$; we have then

$$n_{s+N}(V_0)=\frac{1}{4\tau_0 a_0(\pi)^{\frac{1}{2}}}[\Theta(v_0)-\Theta(v_0-a_0)]$$
$$\times[(\rho\tau_0)\exp(-a_0^2/(\rho\tau_0)^2)$$
$$+a_0(\pi)^{\frac{1}{2}}\Theta(a_0/\rho\tau_0)] \quad \text{(II-3.13)}$$

while the pulse is on.

The most important special case is the rectangular pulse. Here τ_0 and $\tau_0\rho$ both vanish in the limit of finite noise power and vertical pulse front, giving us finally from (II-3.12)

$$n_{s+N}(V_0)=\left\{\frac{\rho}{2\pi}\exp(-(v_0-a_0)^2)\right.$$
$$\left.+\frac{[\Theta(v_0)-\Theta(v_0-a_0)]}{2\tau}\right\}. \quad \text{(II-3.14)}$$

The expected number of peaks exceeding V_0 while the pulse is on is $\tau n_{s+N}(V_0)$, cf. (II-3.14) above.

For a negligible transient contribution we require the pulse width to be related to the noise band width by

$$\tau \gg \frac{\pi}{\rho}\exp(v_0-a_0)^2[\Theta(v_0)-\Theta(v_0-a_0)]. \quad \text{(II-3.15)}$$

Now, the expected number n_{on} of peaks/sec. exceeding V_0 during a single pulse is given by (II-3.14), and when there is no pulse, n_{off} is expressed by (II-3.6). Let T be the repetition period of the pulses. The expected number n_S of *spurious* peaks per second, greater than V_0, is

$$n_S \equiv \{\text{total no. peaks/sec.—no. signal peaks/sec. in absence of noise}\}, \quad \text{(II-3.16)}$$

$$\left.\begin{aligned}&=(\tau/T)n_{on}+(1-\tau/T)n_{off}-1/T\\&=\frac{\rho\exp(-v_0^2)}{2\pi}\left(1+\frac{\tau}{T}\right.\\&\quad\left.\times[\exp(-a_0^2+2v_0a_0)-1]\right)\\&\quad-\alpha(v_0, a_0; T)\end{aligned}\right\} \quad \text{(II-3.17)}$$

where

$$\alpha(v_0, a_0; T)=(1/T)(1-\Theta(v_0)/2$$
$$+\Theta(v_0-a_0)/2). \quad \text{(II-3.18)}$$

Now if the pulse is on only a short fraction of the period $(T \gg \tau)$, which is the customary case in most instances, (II-3.17) simplifies to

$$n_S=\frac{\rho\exp(-v_0^2)}{2\pi}-\alpha(v_0, a_0; t), \quad \text{(II-3.19)}$$

as we would expect, since the number of added noise peaks while the pulse is on is negligible compared with the number exceeding V_0 in the interval $T-\tau$. Note that in the limit of large signals $(a_0^2 \gg v_0^2)$ the transient effect gives α the value $1/2T$, while for small signals $(v_0^2 \gg a_0^2)\alpha$ becomes $1/T$, as then the contribution of the transients can be ignored.

Other pulse shapes, such as the gaussian or "optical," may in principle be handled in the same way, with the aid of (II-3.5), but the integration does not prove readily tractable, which is the pragmatic reason for our choice of the rather idealized pulse shapes treated in the present paper.

4. Peak-Number Distribution Density for Narrow-Band Noise

In this case (see Fig. 1b) our voltage wave $V(t)$ may be represented by (II-3.1) with

$\omega_n = \omega_n' + \omega_0$, where ω_0 is an arbitrary (angular) frequency in the band, usually the central one if the power spectrum is symmetrical, and ω_n' is a "local" frequency, referred to ω_0. Then we have

$$V(t) = V_C(t)\cos\omega_0 t + V_S(t)\sin\omega_0 t. \quad \text{(II-4.1)}$$

Here V_C and V_S are, respectively, the slowly varying parts of the noise wave in phase with $\cos\omega_0 t$ and $\sin\omega_0 t$. To determine the probability density W_1 we require now the joint distribution of the *four* random variables $X_1 = V_C$, $X_2 = V_S$, $X_3 = \dot{V}_C$, $X_4 = \dot{V}_S$. The approach is exactly the same in the case previously considered, of broad-band noise (Section 3), but extended to four variables instead of two. The desired probability density is[5]

$$W_1(X_1X_2X_3X_4) = \frac{1}{4\pi^2(b_0b_2 - b_1^2)} \exp\left\{\frac{-b_2(V_C^2 + V_S^2) - b_0(\dot{V}_C^2 + \dot{V}_S^2) - 2b_1(V_C\dot{V}_S - V_S\dot{V}_C)}{2(b_0b_2 - b_1^2)^2}\right\}, \quad \text{(II-4.2)}$$

where, in analogy with (II-3.4b),

$$b_n \equiv \int_0^\infty (\omega - \omega_0)^n w(f)\,df. \quad \text{(II-4.3)}$$

Now, since the noise is narrow-band, we may introduce the concept of the envelope and write (II-4.1) as

$$V(t) = R(t)\cos[\omega_0 t - \theta(t)]. \quad \text{(II-4.4)}$$

The quantity R is the envelope and θ is a phase, both of which are slowly varying functions of the time compared to $\cos\omega_0 t$. Comparison of the coefficients of $\cos\omega_0 t$ and $\sin\omega_0 t$ in (II-4.4) and (II-4.1) shows at once that

$$R\cos\theta = V_C = X_1; \quad R\sin\theta = V_S = X_2. \quad \text{(II-4.5)}$$

The Jacobian of the transformation from X-space to $(R, \dot{R}; \theta, \dot{\theta})$-space is R^2, so that

$$W_1(X_1X_2X_3X_4) = R^2 w_1(R, \dot{R}; \theta, \dot{\theta}) = W_1(R, \dot{R}; \theta, \dot{\theta}), \quad \text{(II-4.6)}$$

where $w_1(R, \dot{R}; \theta, \dot{\theta})$ is the density obtained on substituting for X_1, X_2, etc., their values in R, θ, etc., obtained from (II-4.5) and its time derivative. To obtain $W_1(R, \dot{R})$, the probability density of the envelope and its rate of change, we must integrate over θ and $\dot{\theta}$, the ranges of which are respectively $(0 - 2\pi)$ and $(-\infty$ to $+\infty)$. From (II-4.2)–(II-4.6) we obtain

$$W_1(R, \dot{R}) = \frac{R}{(2\pi\mu^{33})^{\frac{1}{2}}} \exp(-\mu^{33}\dot{R}^2/2|M|) \cdot \exp[-R^2(\mu^{11}\mu^{33} - [\mu^{14}]^2)/2\mu^{33}|M|], \quad \text{(II-4.7)}$$

where

$$|M| = (b_0b_2 - b_1^2)^2; \quad \mu^{11} = b_2|M|^{\frac{1}{2}}; \quad \mu^{33} = b_0|M|^{\frac{1}{2}}; \quad \mu^{14} = b_1|M|^{\frac{1}{2}}.$$

When (II-4.7) is substituted into (II-1.5) we obtain for the probability that the expected number of envelope peaks exceeding R_0

$$p(R_0) = \frac{|M|R_0}{(2\pi[\mu^{33}]^3)^{\frac{1}{2}}} \exp[-R_0^2(\mu^{11}\mu^{33} - [\mu^{14}]^2)/2\mu^{33}|M|]. \quad \text{(II-4.8)}$$

This simplifies somewhat when we observe that $b_1 \doteq 0$ because of the narrow-band condition and assumed symmetry of the noise spectrum. Our result (II-4.8) reduces to

$$p(R_0) = \left(\frac{b_2}{b_0}\right)^{\frac{1}{2}} \frac{R_0}{(2\pi b_0)^{\frac{1}{2}}} \exp(-R_0^2/2b_0) = \left(\frac{b_2}{2\pi}\right)^{\frac{1}{2}} \times (\text{Rayleigh probability density}) \quad \text{(II-4.9)}$$

as the expected number of crossings per second, or equivalently, the expected number of envelope peaks per second greater than R_0. Again because of the stationary character of the process, $p(R_0)$ is independent of the time.

5. Peak-Number Distribution Density for Narrow-Band Signal and Noise

The results of the preceding section may be extended without difficulty to include the case of a modulated signal. The disturbance in question is now represented by

$$V(t) = V_N(t) + V_s(t), \quad \text{where} \quad V_s(t) = A_0(t)\cos\omega_c t \quad \text{(II-5.1)}$$

and where V_N is specified by (II-4.1). Here we let $\omega_c = \omega_0 + \omega_d$, so that ω_d is the difference be-

tween the central noise (angular) frequency ω_0 and the carrier frequency ω_c. Then it is evident that

$$V(t)=(V_C+\alpha)\cos\omega_0 t+(V_S-\beta)\sin\omega_0 t =R\cos(\omega_0 t-\theta), \quad \text{(II-5.2)}$$

with

$$\alpha\equiv A_0(t)\cos\omega_d t;\quad \beta\equiv A_0(t)\sin\omega_d t;$$
$$R=[(V_C+\alpha)^2+(V_S-\beta)^2]^{\frac{1}{2}};$$
$$\theta=\tan^{-1}\left(\frac{V_S-\beta}{V_C+\alpha}\right), \quad \text{(II-5.3)}$$

the only additional requirements being that $A_0(t)$ and $\dot{A}_0(t)$ do not vary faster than V_C or V_S and that $\omega_d\ll\omega_0, \omega_c$, as otherwise we would no longer be able to consider R as an envelope. As before (Section 3) we may set

$$Y_1=V_C+\alpha=X_1+\alpha;\quad Y_2=V_S-\beta=X_2-\beta;\ \text{etc.},$$

and use (II-4.2)–(II-4.3). The Jacobian of the transformation from **X**-space to $(R, \dot{R}, \theta, \dot{\theta})$-space is again R^2, where now R is given in (II-5.3). We obtain finally the following general expression for $W_1(R, \dot{R}; \theta, \dot{\theta}; t)$:

$$W_1(R,\dot{R};\theta,\dot{\theta};t)=\frac{R^2}{4\pi^2|M|^{\frac{1}{2}}}\exp(-[Af_1(R,\theta,t) +Bf_2(R,\dot{R},\theta,\dot{\theta},t) +Cf_3(R,\dot{R},\theta,\dot{\theta},t)]), \quad \text{(II-5.4)}$$

where

$$A=\mu^{11}/2|M|;\quad B=\mu^{33}/2|M|;\quad C=2\mu^{14}/2|M|, \quad \text{(II-5.5)}$$

and

$$\left.\begin{aligned} f_1&=R^2+\alpha^2+\beta^2-2R(\alpha^2+\beta^2)^{\frac{1}{2}}\cos(\theta+\phi),\\ f_2&=\dot{R}^2+\dot{\theta}^2R^2+\dot{\alpha}^2+\dot{\beta}^2+[2R\dot{\theta}\sin(\theta+\psi) -2\dot{R}\cos(\theta+\psi)](\dot{\alpha}^2+\dot{\beta}^2)^{\frac{1}{2}},\\ f_3&=R^2\dot{\theta}-[R\dot{\theta}\cos(\theta+\phi) +\dot{R}\sin(\theta+\phi)](\alpha^2+\beta^2)^{\frac{1}{2}}+\dot{\alpha}\beta-\dot{\beta}\alpha +R(\dot{\alpha}^2+\dot{\beta}^2)^{\frac{1}{2}}\sin(\theta+\psi),\end{aligned}\right\} \quad \text{(II-5.6)}$$

with the angles ϕ and ψ defined by

$$\cos\phi=\frac{\alpha}{(\alpha^2+\beta^2)^{\frac{1}{2}}},\quad \sin\phi=\frac{\beta}{(\alpha^2+\beta^2)^{\frac{1}{2}}};$$
$$\cos\psi=\frac{\dot{\alpha}}{(\dot{\alpha}^2+\dot{\beta}^2)^{\frac{1}{2}}},\quad \sin\psi=\frac{\dot{\beta}}{(\dot{\alpha}^2+\dot{\beta}^2)^{\frac{1}{2}}}. \quad \text{(II-5.7)}$$

The desired probability $W_1(R, \dot{R})$ requires integration over θ and $\dot{\theta}$; the latter is easily performed, but the former proves excessively complicated in the general case. Hence we introduce the approximation $b_1\doteq 0$, which makes $C\doteq 0$; see (II-5.5). Integrating over $\dot{\theta}$ and expanding the trigonometric exponentials in a triple series of Bessel functions, we may effect the integration over θ with the aid of

$$I_{jlm}=\frac{1}{2\pi}\int_0^{2\pi}\cos j(\theta+\eta)\times\cos l(\theta+\delta)\cos m(\theta+\gamma)d\theta =\frac{1}{2\epsilon_j}\{\delta_{l-m}{}^j\cos[l(\delta-\eta)-m(\gamma-\eta)] +\delta_{l+m}{}^j\cos[l(\delta-\eta)+m(\gamma-\eta)]. \quad \text{(II-5.8)}$$

If we note that $\lambda\equiv\phi-\psi=-\omega_d A_0(t)/\dot{A}_0(t)$, we find at last that the probability density W_1 is

$$W_1(R,\dot{R};t)=\frac{R}{b_0(2\pi b_2)^{\frac{1}{2}}}\exp[-(R^2+A_0(t)^2)/2b_0-(R^2+\dot{A}_0(t)^2/2+\omega_d{}^2A_0(t)^2/2)/2b_2] \times\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\frac{\epsilon_m}{2}I_k\left(\frac{\dot{\alpha}^2+\dot{\beta}^2}{2b_2}\right)I_m\left(\frac{\dot{R}(\dot{\alpha}^2+\dot{\beta}^2)^{\frac{1}{2}}}{b_2}\right)\left\{\epsilon_{2k+m}I_{2k+m}\left(\frac{RA_0}{b_0}\right)\cos(2k+m)\lambda +\epsilon_{|2k-m|}I_{|2k-m|}\left(\frac{RA_0}{b_0}\right)\cos(2k-m)\lambda\right\}. \quad \text{(II-5.9)}$$

When the carrier is tuned to the center of the *IF* filter, i.e., $\omega_0=\omega_c$, then $\omega_d=0$ and Eq. (II-5.9) is altered accordingly, but the simplification is slight. However, for an unmodulated carrier on center ($\omega_c=\omega_0$), considerable reduction is possible, for then only the terms for $m=k=0$

contribute, and we may write at once from (II-5.9)

$$W_1(R,\dot{R})=\frac{R}{b_0(2\pi b_2)^{\frac{1}{2}}}I_0(RA_0/b_0)\exp[-(R^2+A_0^2)/2b_0-\dot{R}^2/2b_2], \quad \text{(II-5.10)}$$

which is independent of the time, since now the process is stationary (for the envelope and phase, but not for the entire wave). For no signal at all ($A_0=0$), Eq. (II-5.9) reduces further to

$$W_1(R,\dot{R})=\frac{R}{b_0(2\pi b_2)^{\frac{1}{2}}}\exp[-R^2/2b_0-\dot{R}^2/2b_2], \quad \text{(II-5.10a)}$$

a simplified version of (II-4.7).

The probability density $p(R_0;t)$ for the expected number of times per second that our wave will exceed R_0 follows at once from (II-2.6) with the aid of Eq. (2). **13.3** of Watson, *Theory of Bessel Functions* (Cambridge University Press, Teddington, England, 1945). We have

$$p(R_0;t)=\rho\frac{R_0}{2\pi b_0}\exp(-[R_0^2+A_0(t)^2]/2b_0+[\dot{A}_0(t)^2+\omega_d^2A_0(t)^2]/4b_2)$$
$$\times\sum_{k=0}^{\infty}\sum_{m=0}^{\infty}\frac{\epsilon_m\Gamma(m+1/2)}{2m!}\left(\frac{\dot{\alpha}^2+\dot{\beta}^2}{2b_2}\right)^{m/2}I_k\left(\frac{\dot{\alpha}^2+\dot{\beta}^2}{4b_2}\right){}_1F_1\left(\frac{m}{2};m+1;-\frac{\dot{\alpha}^2+\dot{\beta}^2}{2b_2}\right)$$
$$\times\left\{\epsilon_{2k+m}I_{2k+m}\left[\frac{R_0A_0(t)}{b_0}\right]\cos(2k+m)\lambda+\epsilon_{2k-m}I_{2k-m}\left[\frac{R_0A_0(t)}{b_0}\right]\cos(2k-m)\lambda\right\}. \quad \text{(II-5.11)}$$

Again, when there is no modulation and $\omega_d=0$, (II-5.11) assumes a particularly simple form, *viz*:

$$p(R_0)=\rho\frac{R_0}{(2\pi b_0)^{\frac{1}{2}}}I_0\left(\frac{R_0A_0}{b_0}\right)\exp[-(R_0^2+A_0^2)/2b_0]$$
$$=\left(\frac{b_2}{2\pi}\right)^{\frac{1}{2}}\times(\text{Rayleigh dist. for envelope } V_s+V_N). \quad \text{(II-5.12)}$$

We may employ (II-5.12) when the carrier is modulated by a rectangular pulse,[6] for the time the pulse is on, and use (II-5.13) below, for the interval of noise alone:

$$p(R_0)=\rho\frac{R_0}{(2\pi b_0)^{\frac{1}{2}}}\exp[-R_0^2/2b_0]. \quad \text{(II-5.13)}$$

Calendar of Meetings

September

20–24 Illumination Engineering Society (National Technical Conference), Boston, Massachusetts
21–24 International Rheological Congress, Scheveningen (The Hague), Netherlands
27–29 Technical Association of the Pulp and Paper Industry, Poland Spring, Maine
27–30 American Society for Professional Geographers, Madison, Wisconsin
29–October 2 Institute of Radio Engineers, Los Angeles, California

October

5–7 American Institute of Electrical Engineers, Washington, D. C.
6–9 Society of Automotive Engineers (National Aeronautic Meeting and Aircraft Engineering Display), Los Angeles, California
7–9 International Scientific Radio Union (American Section) and the Institute of Radio Engineers, Washington, D. C.
12–16 Fifth National Chemical Exposition, Chicago, Illinois
13–15 American Society of Civil Engineers (Fall Meeting), Boston, Massachusetts
13–16 Electrochemical Society, New York, New York
18–22 American Institute of Electrical Engineers (Midwest General Meeting), Milwaukee, Wisconsin
20–22 American Standards Association (Annual Meeting), New York, New York
21–23 Optical Society of America, Detroit, Michigan
24–27 American Institute of Mining and Metallurgical Engineers El Paso, Texas
25–28 Technical Association of the Pulp and Paper Industry, Buffalo, New York
25–29 American Society for Metals, Philadelphia, Pennsylvania
25–29 Society of Motion Picture Engineers, Washington, D. C.
26 Association of Consulting Chemists and Chemical Engineers, New York, New York
27–29 Gas Discharge Conference, Brookhaven National Laboratory, Upton, Long Island
30 American Mathematical Society, New York, New York

November

3–5 American Institute of Electrical Engineers (Southern District Meeting), Birmingham, Alabama
4–5 Society of Automotive Engineers (National Fuels and Lubricants Meeting), Tulsa, Oklahoma
4–6 Acoustical Society of America, Cleveland, Ohio
4–6 National Electronics Conference, Inc. (Annual Technical Forum), Chicago, Illinois
7–10 American Institute of Chemical Engineers, New York, New York
8–10 Institute of Radio Engineers, Rochester, New York
11–13 Geological Society of America (Annual Meeting), New York, York
12 X-Ray and Electron Diffraction Conference (Southeast Section) Detroit, Michigan
13–15 Conference on Applied Spectroscopy, Pittsburgh, Pennsylvania

Frequency Stabilization of Microwave Oscillators by Spectrum Lines. II*

José L. García de Quevedo** and William V. Smith
Department of Physics, Duke University, Durham, North Carolina

(Received February 2, 1948)

The design of both narrow band high stabilization and broad band lower stabilization spectrum line discriminators is developed in a generalized form applicable to various spectrum lines and cavity designs. Comparison with experiment is made for two broad band discriminators utilizing the NH_3 3,3-absorption line at 23,870 megacycles as the frequency determining element. These discriminators used with an amplifier of 2000 gain, reduce drifts by factors of 250 and 1000, respectively, compared to an unstabilized tube.

INTRODUCTION

THIS paper is an extension of the work of the authors and two co-workers[1] on the use of the NH_3 3,3-line at 23,870 megacycles to control the frequency of microwave oscillators. The theoretical analysis has been generalized to facilitate comparison with other spectrum lines and to include a greater variety of possible discriminator designs. A stabilizer design has been achieved that reduces frequency drifts by a factor of 1000 in comparison with an unstabilized tube—i.e., from long period drifts of a few megacycles to long period drifts of a few kilocycles—and means of obtaining still higher stability are discussed. The design of a broad band frequency standard based on microwave spectrum lines is developed. Saturation of the gas by the microwave radiation is found to be a major factor influencing the discriminator design.

CIRCUIT THEORY

The circuit theory developed in I for the admittance of a magic tee whose two symmetric arms are balanced by two equal lengths of wave guide, one length being filled with a low pressure gas with a resonance absorption at ν_0 (I, Fig. 1b), may be reduced to the problem of a magic tee with symmetric arms balanced by two identical cavities, one filled with gas, and both gas and cavities resonant at ν_0. A particularly simple cavity is a shorted wave guide of length l. The impedance of this cavity when filled with gas is found to be (*I*):

$$Z=Z_0\tanh[\alpha_0+j\beta+(\gamma/2)]l,$$

where α_0 represents the skin loss of the guide, $\beta=2\pi/\lambda_g$, λ_g being the wave-length in the guide, and γ is the complex gas attenuation in cm^{-1} (or twice the attenuation in nepers/cm) including both absorption and anomalous dispersion terms.

From the theory of collision broadening,[2]

$$\gamma=\gamma_0\Delta\omega_0/[\Delta\omega_0+j(\omega-\omega_0)]. \quad (1)$$

At resonance $\gamma=\gamma_0$, the maximum value of absorption, while at $\omega-\omega_0=\Delta\omega_0$, the absorption is reduced by a factor of two and the dispersion is a maximum.

Expanding (1) around $\omega-\omega_0$, and $\tanh 2\pi l/\lambda_{g0}=0$, for small values of α_0 and γ, where λ_{g0} is the value of λ_g at ν_0:

$$Z=\frac{\pi Z_0 l}{\lambda_{g0}}\left[\frac{\lambda_{g0}\alpha_0}{\pi}+j\frac{2(\omega-\omega_0)}{\omega_0}+\frac{\Delta\omega_0\gamma_0\lambda_{g0}}{\pi\omega_0}\times\left[\frac{1}{(2\Delta\omega_0/\omega_0)+j2(\omega-\omega_0)/\omega_0}\right]\right], \quad (2)$$

which is of the form

$$Z=Z_c\left[\frac{1}{Q_c}+j\frac{2(\omega-\omega_0)}{\omega_0}+\frac{1}{y_s[(1/Q_s)+j(2(\omega-\omega_0)/\omega_0)]}\right], \quad (3)$$

* The research described in this report was supported by Contract No. W-28-099-ac125 with the Army Air Forces, Watson Laboratories, Air Materiel Command.

** This work is based on a thesis presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate School of Arts and Sciences of Duke University.

[1] W. V. Smith, José L. García de Quevedo, R. L. Carter, and W. S. Bennett, J. App. Phys. **18**, 1112 (1947). This paper will be referred to hereafter as I. A different approach to the problem is presented by W. D. Hershberger and L. E. Norton, RCA Review **IX**, No. 1 (1948).

[2] J. H. Van Vleck and V. F. Weisskopf, Rev. Mod. Phys. **17**, 227 (1945).

where $Z_c=(\pi l Z_0)/\lambda_{g0}$; $Q_c=\pi/(\lambda_{g0}\alpha_0)$; $y_s=(2\pi Q_s)/(\gamma_0\lambda_{g0})$; $Q_s=\omega_0/(2\Delta\omega_0)$. Thus the shorted transmission line may be thought of as a series resonant cavity of parameters Z_c and Q_c determined by the properties of the cavity without gas, in series with a parallel resonant cavity of parameters $Y_c y_s$ and Q_s where y_s and Q_s are functions of the gas only. The fact that the gas is a parallel circuit while the cavity is a series circuit is a consequence of the anomalous dispersion properties of the gas near resonance and means that the rate of change of impedance with frequency caused by the gas is in the opposite sense to that caused by the cavity alone. Thus filling a cavity of resonant frequency ν_0 with a gas of the same resonant frequency will reduce rather than agument the effects of the cavity.

If, however, the impedance of an empty cavity Z' is subtracted from (3), by means of a magic tee as shown in I, the input admittance to the tee becomes

$$Y=1/(Z-Z')=Y_c y_s[(1/Q_s)+j(2(\omega-\omega_0)/\omega_0)]$$
$$=[(1/\alpha)+j(a/\alpha)]Y_0. \quad (4)$$

In the special case where the cavities are shorted transmission lines, with arms of lengths l and $l+\lambda_g/8$, a discriminator curve may be obtained as in I, giving a slope of

$$(dV/d\nu)|_{\nu=\nu_0}=DP_0[Q_s/\nu_0][2\alpha/(1+\alpha)^2] \quad \text{I(7)}$$

volts per megacycle, where P_0 is the input power to the discriminator and D is the rectification efficiency of the crystal. More generally, the magic tee plus two cavities may be treated as a single cavity and then used as in I, circuit 1a, in which case Eq. I(7) is still valid.

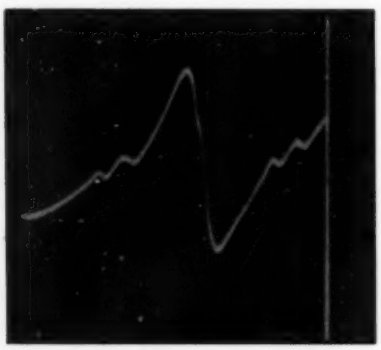


FIG. 1. Discriminator characteristic of 2.16 meter cell using the NH_3 3,3-line. Vertical line establishes axis. The spacing between the center of the linear portion of the discriminator curve and the first satellite is 1.72 megacycles, establishing the frequency scale.

By adjusting the coupling to the cavity, maximum slope is obtained for $\alpha=1$, in which case the gas is equivalent to a cavity of unloaded $Q=Q_s$. If cavity dimensions are greater than the mean free path in the gas and if the residual doppler line width is not approached too closely, $Q_s=\omega_0/(2\Delta\omega_0)$ for the NH_3 3,3-line may be made equal to 23,870/0.2 or about 100,000 with the attenuation still given by (1) and the discriminator slope by I(7). There is one important restriction, however. The r-f fields in the cavity must be low enough to avoid saturation, which implies either a large cavity or a low power input, as will be discussed in the following section.

SATURATION EFFECT

When high powers are passed through a gas at low pressure, it is found that the thermal equilibrium of energy states is altered, with a resultant change in the absorption at the resonant frequency and in the line width, so that[3]

$$(\gamma)_{\nu_0}=\frac{\gamma_0}{[1+(AP/(\Delta\nu_0)^2)]},$$
$$2\Delta\nu=2\Delta\nu_0[1+(AP/(\Delta\nu_0)^2)]^{\frac{1}{2}}, \quad (5)$$

where P is the incident power in ergs/cm^2 monochromatic at the frequency in question,

$$A=8\pi|\mu_{ij}|^2/3ch^2.$$

μ_{ij} = matrix element of the dipole moment; $\gamma_0=((\pi h^2\nu^2)/(kT\Delta\nu_0))n_0A$; n_0 = number of molecules/cc in ground state. For the NH_3 3,3-line, A is 1×10^7.*** While relations (5) were derived for plane traveling waves, they are very nearly true for traveling waves in wave guide if P is replaced by the $\sqrt{2}$ times the average power and both γ_0 and A are multiplied by λ_g/λ. For standing waves a second averaging over the guide length replaces P by twice the average power, or four times the incident power.

[3] R. L. Carter and W. V. Smith, Phys. Rev. **73**, 1053 (1948); also C. H. Townes, Phys. Rev. **70**, 665 (1946); B. Bleaney and R. P. Penrose, Proc. Phys. Soc. **60**, 83 (1948); R. Karplus and J. Schwinger, Phys. Rev. **73**, 1020 (1948); R. Karplus, Phys. Rev. **74**, 223 (1948).

*** Appropriate averaging (see reference 3) over the Zeeman components of the line replaces A by 1.7×10^7.

If the coupling to the balanced cavities is adjusted to give $\alpha=1$, all power incident at the resonant frequency is adsorbed. Assuming the gas absorption to be large compared to the skin losses, as is the case for the NH_3 3,3-line and most silver cavities operating in high modes, then the Q_s of the gas may be related to the cavity volume V, the cross section ab of guide feeding the cavity, and the ratio k of energy density in cavity and guide as follows:

$$Q_s=\frac{2\pi\times\text{stored energy}}{\text{energy dissipated per cycle}}$$

$$=\frac{\epsilon_0\langle E_{\text{cav}}^2\rangle_{\text{Av}}V\omega}{2ab\bar{P}_{\text{guide}}}=\frac{\pi}{2}\frac{\lambda_{g0}}{\lambda^2}\frac{Vk}{ab} \quad (6)$$

in m.k.s. units. (In the special case where the cavities consist of sections of wave guide of the same cross section as the tee, $k=\frac{1}{2}$ from the power division at the tee.)

Substituting (5) in I(7), for $\alpha=1$:

$$(dV/d\nu)|_{\nu_0}=\frac{DP_0}{4\Delta\nu_0\left[1+\dfrac{4AP_0k}{ab(\Delta\nu_0)^2}\dfrac{\lambda_{g0}}{\lambda}\right]^{\frac{1}{2}}}. \quad (7)$$

For fixed input powers, relation (7) increases as $\Delta\nu_0$ decreases irrespective of the value of k, although it increases more rapidly as k approaches zero, corresponding to a large cavity, low field in the cavity, and absence of saturation. If a small cavity is employed and the energy density within the cavity is large enough to cause saturation, the gas attenuation soon becomes smaller than the copper losses and the above analysis no longer holds.

While for a given volume the steepest discriminator curves are obtained when the coupling is adjusted to get $\alpha=1$, for many purposes the broad band discriminator design employing balanced wave guide arms is desirable. By using tapered sections from the tee to arms employing larger sized guide, k may be made less than $\frac{1}{2}$ while the broad band feature of the discriminator is preserved. For a given volume and input power however, α is no longer a variable. Assuming α to be much smaller than 1, substitution in I(7) then gives

$$(dV/d\nu)|_{\nu_0}=DP_0\alpha/\Delta\nu=K\gamma_0P_0/\Delta\nu, \quad (8)$$

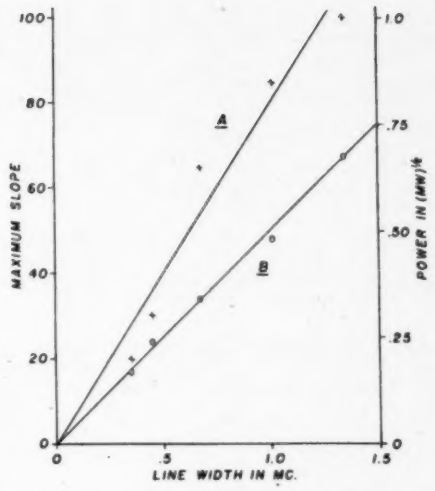


FIG. 2. Maximum discriminator slope and corresponding input power for 2.16-meter cell as function of line width. Curve A: observed maximum discriminator slope. Curve B: input power for maximum slope.

where $K=(D\lambda_{g0}Z_c)/(2\pi Z_0)$. Differentiating (8) with respect to $\Delta\nu_0$ and equating the resulting expression to zero for a maximum gives

$$4P_0kA\lambda_{g0}/ab(\Delta\nu_0)^2\lambda=\tfrac{1}{2}. \quad (9)$$

Converting units from c.g.s. to P in milliwatts and line width $2\Delta\nu_0$ in megacycles, with $ab=1.07\times0.43$, and $\lambda_g/\lambda=1.2$:

$$P_0/(2\Delta\nu_0)^2=0.115/kA\times10^{-7}. \quad (10)$$

Since the maximum value of discriminator slope occurs for $AP/(\Delta\nu_0)^2=\frac{1}{2}$, where saturation is just beginning, the ratio of signal to crystal noise is also near its optimum value.

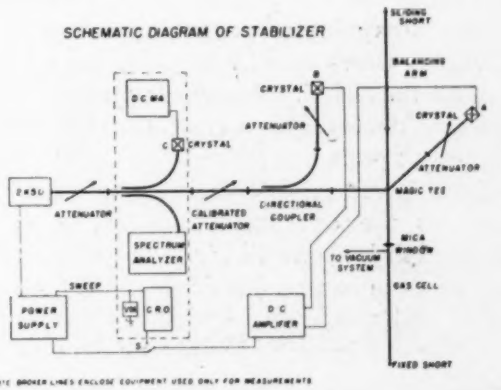


FIG. 3. Schematic diagram of the laboratory model of the spectrum line stabilizer. The switch s is shown in the "stabilized" position. The cathode-ray oscilloscope was used to observe the discriminator curve, and the spectrum analyzer was used to measure the stabilization factor.

TABLE I.

Designation	Frequency[a] (megacycles)	Strength[b] relative to 3,3
2,2	23,722.63	0.4
3,3	23,870.13	1.0
4,4	24,139.41	0.6
5,5	24,532.98	0.5
6,6	25,056.02	0.9

[a] W. E. Good and D. K. Cole, Phys. Rev. **71**, 383 (1947).
[b] James W. Simmons and Walter Gordy, Phys. Rev. **73**, 713 (1948); also C. H. Townes, Phys. Rev. **70**, 433 (1947).

In the experiments to be discussed, the cavities were shorted wave guides of varying dimensions connected to the arms of the tee by tapers, in which case substitution of (10) in (8) yields

$$(dV/d\nu)|_{\nu_0} = 5\times10^{-2}(DV\gamma_0/ab) \times[P_0k/(A\times10^{-7})]^{\frac{1}{2}} \quad (11)$$

volts per megacycle, where the crystal rectification efficiency D is in volts per milliwatt and V is the volume of the cavity. The independent variable is chosen as P_0 as practical considerations, such as crystal saturation or tube pulling, limit its maximum value to about 10^{-1} mw in the 20,000 to 30,000 megacycle region. From the definitions of γ_0 and A, the maximum discriminator slope is seen to be proportional to the fraction of molecules in the ground state and to the effective dipole moment of the transition.

COMPARISON OF EXPERIMENT WITH THEORY

A discriminator was constructed using balance arms of 2.16 meters and $2.16+\lambda_g/8$ meters of 1.07×0.43 cm silver wave guide. Photographs of the discriminator output as a function of frequency were taken by sweeping the oscillator through the NH_3 3,3-resonant frequency. Several series of photographs were taken for different fixed input powers to the tee, and a continuously variable pressure (proportional to line width $2\Delta\nu_0$) as the cell was evacuated. A typical photograph is shown in Fig. 1 in which the fine structure can be seen (establishing the frequency scale[4]) and a reference line shows the vertical axis. The maximum and minimum discriminator outputs occur at the half-power points, so that their separation is the line width $2\Delta\nu_0$. Also, the maximum discriminator slope at the resonant frequency can be shown to be twice the ratio of the difference of maximum and minimum discriminator outputs to the line width. The vertical scale was established by a separate calibrating signal not shown on the photograph.

As the cell was evacuated, the maximum discriminator output continually decreased, at first very slowly, corresponding to the region of approximately constant γ_0, and then more rapidly as saturation was approached. During this time the line width decreased, so that the discriminator slope first increased, then went through a maximum as saturation was approached, finally decreasing again at the lowest pressures. A plot of maximum discriminator slope as a function of line width is shown in Fig. 2, curve A being the observed points. Also plotted is the square root of the power at which the maximum slope occurred. Both relations are linear, as predicted theoretically. The plot of the square root of the power *versus* the line width yields an experimental value of $P_0/(2\Delta\nu_0)^2=0.25$, which from Eq. (10) gives a value of A of 0.97×10^7. The best theoretical and experimental value of A from other investigations[2] is 1.7×10^7. The discrepancy can at least in part be ascribed to the more complicated averaging of electric fields necessary in the discriminator investigation (an averaging that was done only semiquantitatively). The maximum observed slope was 34 millivolts per megacycle. The decrease of a factor of two from the figure of 70 millivolts per megacycle quoted in I is because the present figure is for one crystal only, whereas the previous higher value was for the output of two balanced crystals.

A second cell of dimensions $1.07\times2.40\times386$ cm was also tested, and a maximum slope of 152 millivolts per megacycle for a single crystal was observed. Theoretically, this large cell should give a slope steeper than the small one by the ratio $((386)/(216))(2.40/0.47)^{\frac{1}{2}}=4.22$. Different crystals were used in the two cases however, and the observed improvement with the large cell was somewhat greater than anticipated.

FREQUENCY STANDARD DESIGN

The small cell discriminator discussed in the previous section was built into a frequency

[4] B. P. Dailey, R. L. Kyhl, M. W. P. Strandberg, J. H. Van Vleck, and E. B. Wilson, Jr., Phys. Rev. **70**, 984 (1946).

stabilizer according to the design of reference I, except that the correction signal was adjusted to be zero at the resonant frequency of the gas by balancing the output of the magic tee against the output of a directional coupler instead of a second magic tee (see Fig. 3). The method yielded a much smaller background of reflections than did the two-magic-tee set-up previously used, but resulted in a loss of a factor of two in sensitivity.

The stabilization was measured by modulating the reflector of the 2K50 both when stabilized and when unstabilized, and comparing the frequency shift per unit voltage in the two cases. With the measured d.c. amplifier gain of 2000 and tube tuning characteristic of 283 millivolts per megacycle, the observed maximum discriminator slope of 34 millivolts per megacycle should give a stabilization factor of 240. The observed stabilization factor was 125, which is in satisfactory agreement since only half of the output signal of the amplifier is available for stabilization. Typical fluctuations in power supply voltage of 0.1 volt applied to the reflector correspond to frequency fluctuations of about 1.5 kilocycles. A stabilizer based on the large cell would have a stability approximately fourfold greater while, by using appropriately large cavities and adjusting the coupling for maximum results, much greater stability can be achieved in less flexible systems. It should be emphasized that higher gain amplifiers can also be used and that the modifications in the r-f circuit necessary to use an i-f amplifier[5] do not affect the basic discriminator design.

In applying the theoretical analysis of the preceding section to the design of practical frequency standards, it is evident that the requirements of maximum stabilization and flexibility of operation are to some extent mutually incompatible. Thus, a practical broad band system would have a cell of 10.7×4.3 mm cross section to match the tee cross section, and a length of about 6 meters (which can, of course, be coiled compactly), as beyond this length the sharpness of the discriminator curve is reduced by departure of $\tanh(\alpha_0+\gamma_0/2)l$ from linearity with $(\alpha_0+\gamma_0/2)l$, at least for a silver cell and the NH_3 3,3-line. The discriminator response is not critical to the eighth wave-length difference in balance arm length, but some retuning is necessary for different lines. Some of the frequencies and discriminator characteristics available for ammonia are cited in Table I. To get the discriminator slope in millivolts per megacycle for the above cell, multiply column three by 100. The table is for the stronger lines only, and at room temperature. If the cell is immersed in dry ice, the density of molecules is greater for a given line width by $(300/190)^{\frac{1}{2}}=1.25$, thus increasing γ_0 and reducing saturation effects, while the relative intensities of the low J number lines is increased over that of the high J number lines. The ultimate limiting factor of the doppler width is also reduced at low temperatures.

To date, ammonia is the only gas with an absorption in the centimeter region strong enough to be practical in a broad band frequency standard. It should be noted, however, that many transitions in the millimeter region are even stronger than the ammonia lines.

The use of a cell of large cross section connected by a taper to the magic tee is an attempt to gain sensitivity while preserving the broad band feature of the design. This can indeed be done, but it should be noted that for a given cell length, the gain in sensitivity is proportional to the square root of the cross section when coupling is by means of a taper, as for a constant input power the field strength in the guide varies inversely with the cross section, and the optimum line width therefore inversely with the square root of the cross section. If size is no objection, however, large cross-section cells can be used in broad banded systems. Some trouble with extraneous modes of propagation was observed in the large cell discussed previously, but these could always be tuned out by squeezing the cell walls to slightly detune the resonance.

For the most precise control at one fixed frequency, it is necessary, of course, to maintain the balance arms at the same temperature, although it is not necessary that the temperature remain constant. A more compact design, capable of better temperature control, might be accomplished by using high Q cylindrical cavities of minimum surface to volume ratio, and coupling

[5] R. V. Pound, Rev. Sci. Inst. **17**, 490 (1946).

to them rather heavily so that the reflection coefficient at resonance is 0.9 or more and a slight unbalance of the two cavities would be unimportant compared to the high Q gas absorption. With this reflection coefficient and short sections between tee and cavity, a uniform but not necessarily constant temperature would again be necessary for high stability.

In conclusion, the authors wish to thank Mr. Casimir Nawrocki for his invaluable assistance in pointing out the peculiar adaptability of the 2K50 oscillator to this problem, and in helping familiarize us with the suitable operation of the tube. We also wish to thank Dr. W. O. Gordy and Dr. W. M. Nielsen for their continuing interest in the problem.

On the Calculation of Radiation Patterns of Dielectric Rods*

R. B. Watson and C. W. Horton
Defense Research Laboratory, The University of Texas, Austin, Texas
(Received February 13, 1948)

A previous paper described the radiation field of a dielectric rod used as a termination for a wave guide. The theory therein made use of a standing wave distribution of equivalent currents over the surface of the rod to describe its radiation effect. It is shown in this paper that the equivalent currents may be expressed in terms of traveling waves without appreciably changing the computed radiation patterns.

THE theory for the radiation field of a dielectric rod as a termination for a wave guide has been developed and a summary of results presented in the paper, "The radiation patterns of dielectric rods—experiment and theory."[1] In this paper certain equivalent currents distributed over the surface of the dielectric rod are used to describe its radiation effect. In the way in which the theory is there formulated, these distributions of currents take the form of standing waves over the surfaces of the rod. Since there is appreciable radiation occurring, it might be more appropriate to approximate these currents by expressions representative of traveling waves. The surface magnetic currents described in Eq. (9) of the above-mentioned paper should then take the following form:

$$\mathbf{M}=\mathrm{j}M_0\cos\frac{\pi y}{b}e^{j(\omega t-k'z)}. \tag{1}$$

The analysis proceeds as before except that the integral I_2 [Eq. (19)] has the form

$$I_2=\int_0^l e^{jk'z}e^{-jkz\cos\theta}dz, \tag{2}$$

which can be readily integrated to give

$$I_2=2\left\{\frac{\sin qkl}{q}+j\frac{1-\cos qkl}{q}\right\} \tag{3}$$

or

$$|I_2|^2=\frac{16}{q^2}\sin^2\frac{qkl}{2}, \tag{4}$$

where

$$q=n-\cos\theta, \quad n=k'/k.$$

This is a much simpler expression than the prior value for I_2, and the maxima and the minima of the absolute value are readily found by analysis. For the values of $n=1.10$, $kl=26.4$, the function is found to have minima at $\theta=30.5°$, 51.5°, 67.3°, etc. The function has maxima at $\theta=0°$, 40.6°, 59°, etc. The corresponding values for the electric field for the first three of these maxima are 0 db, −12.5 db, and −18 db. Comparison with Fig. 4 of the reference shows that the differences of these values from the values there plotted are insignificant. The terms other than I_2 in the

* The work described in this paper was done at the Defense Research Laboratory under the sponsorship of the Bureau of Ordnance, Navy Department, Contract NOrd-9195.

[1] R. B. Watson and C. W. Horton, J. App. Phys. **19**, 661 (1948).

calculation of the field remain the same, and therefore it is concluded that the traveling wave picture does not appreciably modify the radiation pattern as compared to the standing wave picture. The calculation for $kl=64.2$, because of the periodicity of the function $\sin^2(qkl/2)$, leads to a very small value for the maximum at 0° and a very large value, of the order of 20 db greater, for the next maximum which occurs at 16.2°. This is in agreement with the calculation for the standing wave type of currents.

A similar evaluation may be carried out for the equivalent electric currents on the sides of the dielectric rod. When this is done the integral I_2 takes the form of Eq. (2), and the evaluation proceeds in the same way. This process has already been shown to produce the same maxima and minima as the integral for the standing wave description, and hence the use of traveling waves for the description of the electric currents produces no new results.

It thus appears that neither a standing wave description nor a traveling wave description of the equivalent currents on the surface of the dielectric rod is satisfactory for describing the radiation patterns for long rods, that is, greater than about six wave-lengths, but that either description gives agreement with experimental patterns for rods in the range of length of three to six wave-lengths.

Heat Transfer from a Rotating Disk to Ambient Air

CARL WAGNER
Ordnance Research and Development Division, Sub-Office Rocket, Fort Bliss, Texas
(Received February 19, 1948)

The coefficient of heat transfer from a rotating disk to ambient air in case of a laminar boundary layer is theoretically calculated on the basis of von Kármán's calculation of flow conditions.

INTRODUCTION

NEAR the surface of a rotating disk, air flows mainly in the tangential direction of rotation. An additional component of airflow in radial direction is caused by the centrifugal forces due to rotation. Accordingly, ambient air moves from the surroundings towards the surface of the disk. Calculation of flow velocity as function of the spatial coordinates and the angular velocity has been carried out by Theodore von Kármán[1] in order to obtain the torque momentum due to friction. On the basis of these calculations, the coefficient of heat transfer from a rotating disk to ambient air in case of a laminar boundary layer can readily be obtained as shown below.

MATHEMATICAL ANALYSIS

According to von Kármán, the velocity components, v_r and v_t, of the air flowing in the radial and the tangential direction inside the boundary layer are approximately*

$$v_r=r\omega\cdot\left[1.026\cdot\frac{x}{\delta}\cdot\left(1-\frac{x}{\delta}\right)^2\cdot\left(1+2\frac{x}{\delta}\right)-\frac{1}{2}\cdot\left(\frac{x}{\delta}\right)^2\cdot\left(1-\frac{x}{\delta}\right)^2\right]\quad\text{if}\quad x\leqq\delta \tag{1}$$

$$v_t=r\omega\cdot\frac{1}{2}\cdot\left(2+\frac{x}{\delta}\right)\cdot\left(1-\frac{x}{\delta}\right)^2\quad\text{if}\quad x\leqq\delta. \tag{2}$$

Here r represents the distance from the axis, x the distance from the disk, ω the angular velocity, and δ the thickness of the boundary layer which is independent of r and is given by the formula

$$\delta=2.58\cdot(\nu/\omega)^{\frac{1}{2}} \tag{3}$$

where ν is the kinematic viscosity of air. Equations (1) and (2) hold true inside the boundary layer, whereas outside the boundary layer the

[1] Theodore von Kármán, Zeits. f. angew. Math. u. Mech., 1, 233 (1921); NACA Technical Mem. No. 1092 (1946).

* In Eq. (2) of the present paper an erratum in Kármán's Eq. (27) is corrected; the term $(1-x/\delta)$ must be employed in the second power instead of the first power, as can be derived from von Kármán's subsequent equations.

velocity components v_r and v_t vanish according to the above approximations.

The velocity component v_x, normal to the surface of the disk, is not explicitly given in von Kármán's paper, but can be obtained from a differential equation. Thus, inside the boundary layer

$$v_x = -\int_0^x 2\cdot\frac{v_r}{r}\cdot dx$$

$$= -2\omega\cdot\int_0^x\left[1.026\cdot\frac{x}{\delta}\cdot\left(1-\frac{x}{\delta}\right)^2\cdot\left(1+2\frac{x}{\delta}\right)\right.$$

$$\left.-\frac{1}{2}\cdot\left(\frac{x}{\delta}\right)^2\cdot\left(1-\frac{x}{\delta}\right)^2\right]\cdot dx \quad \text{if} \quad x\leqq\delta, \quad (4)$$

whereas outside the boundary layer the velocity component v_x is constant, *viz.*,

$$v_x = -\int_0^\delta 2(v_r/r)\cdot dx = -0.708\cdot(\nu\omega)^{\frac{1}{2}} \text{ if } x>\delta. \quad (5)$$

At a steady state the general equation of heat balance for a fluid in motion is

$$k\cdot\text{div grad}\vartheta - \text{grad}(\rho\mathbf{v}c_p\vartheta) = 0, \quad (6)$$

where ϑ is the difference between the local temperature and the bulk temperature of air, k the heat conductivity, c_p the specific heat at constant pressure, ρ the density, and $\mathbf{v}$ the vector of flow velocity. The first term in Eq. (6) represents the increase of enthalpy due to heat conduction in the case of air at rest. The second term equals the gradient of the flux of enthalpy irrespective of heat conduction, given by the product of the vector $\rho\mathbf{v}$ of mass flow and the excess of enthalpy $c_p\vartheta$ per unit mass. In cylindrical coordinates, there is no dependency on the angular coordinate because of symmetry. Thus,

$$\text{div grad}\,\vartheta = \frac{\partial^2\vartheta}{\partial r^2}+\frac{1}{r}\cdot\frac{\partial\vartheta}{\partial r}+\frac{\partial^2\vartheta}{\partial x^2}. \quad (7)$$

At a steady state the divergence of the product $\rho\mathbf{v}$ vanishes according to the condition of continuity of mass flow. Consequently the second term of Eq. (6) becomes

$$\text{grad}(\rho\mathbf{v}c_p\vartheta) = c_p\vartheta\cdot\text{div}\mathbf{v}\rho+\mathbf{v}\rho c_p\cdot\text{grad}\vartheta$$

$$=\mathbf{v}\rho c_p\cdot\text{grad}\vartheta=\rho c_p\cdot\left(v_r\cdot\frac{\partial\vartheta}{\partial r}+v_x\cdot\frac{\partial\vartheta}{\partial x}\right). \quad (8)$$

On introduction of expressions (7) and (8) into Eq. (6) and division by $c_p\rho$, it follows that

$$a\cdot\left(\frac{\partial^2\vartheta}{\partial r^2}+\frac{1}{r}\cdot\frac{\partial\vartheta}{\partial r}+\frac{\partial^2\vartheta}{\partial x^2}\right)-v_r\cdot\frac{\partial\vartheta}{\partial r}-v_x\cdot\frac{\partial\vartheta}{\partial x}=0 \quad (9)$$

where a represents the thermal diffusivity of air, given by

$$a = k/c_p\rho. \quad (10)$$

The over-temperature ϑ_0 of the disk shall be given so that the boundary conditions pertaining to differential Eq. (9) are

$$\vartheta=\vartheta_0 \quad \text{at} \quad x=0 \quad (11)$$

$$\vartheta=0 \quad \text{at} \quad x=\infty. \quad (12)$$

According to von Kármán, the thickness of the boundary layer is independent of the distance r from the axis and the radial flow velocity at a given distance x from the disk is proportional to r. Consequently the total mass rate of air flowing in the radial direction at any distance r with the circumference $2\pi r$ is proportional to the area πr^2 of the circle of radius r. Thus, a temperature distribution dependent on x but independent of r can be assumed, so the partial differential Eq. (9) becomes an ordinary differential equation

$$a(d^2\vartheta/dx^2)-v_x(d\vartheta/dx)=0. \quad (13)$$

Upon first integration,

$$(d\vartheta/dx) = -\exp\left[(1/a)\int_0^x v_x dx+A\right] \quad (14)$$

and subsequently, upon second integration,

$$\vartheta = \int_x^\infty \exp\left[(1/a)\int_0^x v_x dx+A\right]\cdot dx. \quad (15)$$

Here A is an integration constant to be calculated from boundary condition (11); the second boundary condition (12) is automatically satisfied by the upper limit of the first integral of Eq. (15). Introduction of Eq. (15) into (11) gives

$$\vartheta_0 = \int_0^\infty \exp\left[(1/a)\int_0^x v_x dx+A\right]\cdot dx. \quad (16)$$

On division of Eqs. (14) and (15) by (16), the integration constant A drops out. Thus

$$\frac{d\vartheta}{dx} = -\vartheta_0 \cdot \frac{\exp\left[(1/a)\int_0^x v_x dx\right]}{\int_0^\infty \exp\left[(1/a)\int_0^x v_x dx\right]\cdot dx} \tag{17}$$

$$\vartheta = \vartheta_0 \cdot \frac{\int_x^\infty \exp\left[(1/a)\int_0^x v_x dx\right]\cdot dx}{\int_0^\infty \exp\left[(1/a)\int_0^x v_x dx\right]\cdot dx}. \tag{18}$$

On the basis of Eq. (17), the heat transfer coefficient h, defined as the quotient of heat flux by temperature difference ϑ_0, is

$$h = \frac{1}{\vartheta_0}\cdot\left[-k\cdot\left(\frac{\partial\vartheta}{\partial x}\right)_{x=0}\right] = \frac{k}{\int_0^\infty \exp\left[(1/a)\int_0^x v_x dx\right]\cdot dx}. \tag{19}$$

The integral involved in Eq. (19). where v_x is defined by Eqs. (4) and (5), is to be calculated by numerical integration. Thus

$$h = 0.339\cdot k\cdot(\omega/\nu)^{\frac{1}{2}} \tag{20}$$

for $\nu/a = 0.74$, valid for air.

In analogy to von Kármán's calculation of the torque momentum, the heat transfer coefficient can also be calculated by means of the following consideration. The enthalpy flow from the disk of surface πr^2 to the boundary layer must equal the enthalpy flow in the boundary layer in the radial direction at the distance r from the axis. Thus

$$\pi r^2 h\vartheta_0 = \int_0^\delta \rho v_r\cdot\vartheta c_p\cdot 2\pi r\cdot dx \tag{21}$$

and consequently, on division by $\pi r^2\vartheta_0$ and introduction of Eqs. (18) and (10),

$$h = \frac{2k}{a}\frac{\int_0^\delta\left\{(v_r/r)\cdot\int_0^\infty \exp\left[\left(\frac{1}{a}\right)\cdot\int_0^x v_x dx\right]\cdot dx\right\}\cdot dx}{\int_0^\infty \exp\left[(1/a)\cdot\int_0^x v_x dx\right]\cdot dx} \tag{22}$$

where the velocity components v_r and v_x are given by Eqs. (1) and (4). Numerical evaluation yields

$$h = 0.335\cdot k\cdot(\omega/\nu)^{\frac{1}{2}} \tag{23}$$

for $\nu/a = 0.74$, valid for air.

The method employed for the result in Eq. (23) is to be regarded as a closer approximation than that employed for Eq. (20), for representations of velocity components in Eqs. (1) and (2) are only approximations, and small deviations from the true velocity functions cause a greater error in differentiation of the function $\vartheta(x)$ as applied in (19) than in integration as applied in (21). However, the difference between the resulting Eqs. (20) and (23) is insignificant.

Equation (23) is expected to hold true as long as the flow in the boundary layer is essentially laminar. According to friction measurements by W. Schmidt,[2] this condition is met if the Reynolds number,

$$Re = (4\pi r^2\omega/\nu), \tag{24}$$

is less than 500,000.

[2] W. Schmidt, Zeits. d. Vereins Deuts. Ing. **65**, 441 (1921).

An Interferometric Method for the Calibration of Electron Microscope Magnification

J. L. Farrant and A. J. Hodge
Division of Industrial Chemistry, Council for Scientific and Industrial Research, Melbourne, Australia
(Received February 6, 1948)

Interferometric methods have been used to measure uniform cylindrical silica and glass fibers as small as 5μ in diameter with better than 1 percent accuracy. Such fibers yield sharp images when examined in the electron microscope and constitute a convenient and rugged specimen for accurate direct calibration of the instrument for magnifications up to 10,000 diameters. Higher magnification ranges are then readily determined by means of an internal calibration obtained by taking micrographs of glass spheres about 0.5 micron in diameter. Such spheres are also convenient specimens for determining the image distortions which must be known to correct the widths of the fiber images. This method appears to be capable of greater accuracy than any employed hitherto.

INTRODUCTION

THE increasing use being made of the electron microscope for the measurement of long spacings in fibrous materials[1,2] and of the dimensions of macromolecules[3,4] and the desirability of correlating these results with those obtained by x-ray diffraction and ultra-centrifuge methods, respectively, have created a demand for accurate means of electron microscope calibration.

Four different calibration methods have been described in the literature. Borries and Ruska[5] used a "two-step" procedure in which the magnifications of the first and second stages are determined separately by imaging suitable apertures with each stage individually. The aperture diameters are large enough to be measured accurately with a light microscope. Another "two-step" method was developed by the Columbian Carbon Company[6] in which a specimen is moved through known distances by means of a micrometer screw. A number of workers have used a direct comparison method[7] in which micrographs of the same specimen are taken with both the electron microscope and a calibrated light microscope. Burton, Barnes, and Rochow[7] have developed a method in which a thin replica of a diffraction grating of known periodicity is used as an electron microscope specimen.

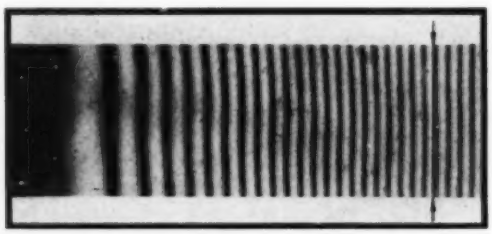


FIG. 1. Fiber No. 10 between a cover glass and an optical flat illuminated with the Hg λ5461 line.

The use of the two-step methods mentioned has lapsed as the errors involved amount to about 20 percent.

The direct comparison method is useful, particularly at magnifications up to 1000 diameters, but is limited in accuracy for two reasons. The light microscope when used without oil immersion has a limit of resolution of 0.25 micron corresponding to a useful magnification of 750 diameters, so that a single particle must be at least 25 microns in diameter to enable its size to be measured with an accuracy better than 1 percent by means of a light microscope. Such a particle fills the 5-cm field of the RCA electron microscopes at a magnification of 2000 diameters; so the higher ranges up to 20,000 diameters must be calibrated by means of internally deter-

[1] C. E. Hall, M. A. Jakus, and F. O. Schmitt, J. Am. Chem. Soc. **64**, 1234 (1942); J. App. Phys. **16**, 459 (1945).
[2] J. L. Farrant, E. H. Mercer, and A. L. G. Rees, Nature **159**, 535 (1947).
[3] W. C. Price, R. C. Williams, and R. W. G. Wyckoff, Arch. Biochem. **9**, 175 (1946).
[4] W. C. Price and R. W. G. Wyckoff, Nature **157**, 764 (1946).
[5] B. v. Borries and E. Ruska, Naturwiss. **27**, 577 (1939).
[6] Columbian Carbon Company: "The particle size and shape of colloidal carbon as revealed by the electron microscope," Binney and Smith Company, Distributors, New York (1940).
[7] C. J. Burton, R. B. Barnes, and T. G. Rochow, Ind. Eng. Chem. **34**, 1429 (1942).

mined ratios of the electron microscope magnifications. This difficulty can be avoided by using a specimen consisting of individual isometric particles about 0.2 micron in diameter, separated by distances of the order of 5–10 microns and supported by a thin membrane. The second objection, that the supporting membrane may temporarily stretch or contract while under the influence of the electron beam, is then encountered.

The fourth method is that most generally used. It has been found[7] that the periodicity of the replica is rarely the same as that of the grating from which it is prepared. It is usual to attempt to avoid this by using either silica replicas prepared by the method of Heidenreich and Peck[8] or metal replicas as suggested by Wyckoff and Williams.[9] The periodicity of the replica may be checked by mounting it on a spectrometer and determining the angles through which it diffracts light of known wave-lengths. This procedure, however, cannot be applied to determine any permanent or temporary local irregularity of the spacings over areas as small as those encountered in the electron microscope at high magnifications. Such permanent irregularities may appear owing to local strains during preparation of the replica or as a result of variations or periodic errors in the rulings. We have noticed temporary variations in the spacings of the replicas while under the influence of the electron beam in the case of collodion and Formvar replicas even when "shadow cast" with up to 30A thickness of gold.

An alternative to these methods would be to employ as a calibration specimen a uniform cylindrical fiber a few microns in diameter. Such fibers are easily drawn from glass or silica. Fine quality glass cloth is a convenient source of 5–6 micron fibers.

However, the accurate determination of the diameter of fibers less than 10 microns in diameter presents some difficulty. Burton, Barnes, and Rochow[7] suggested measuring the diameter by weighing on a microbalance, but rejected the proposal because fibers cannot be obtained uniform over the considerable length required to give sufficient weight.

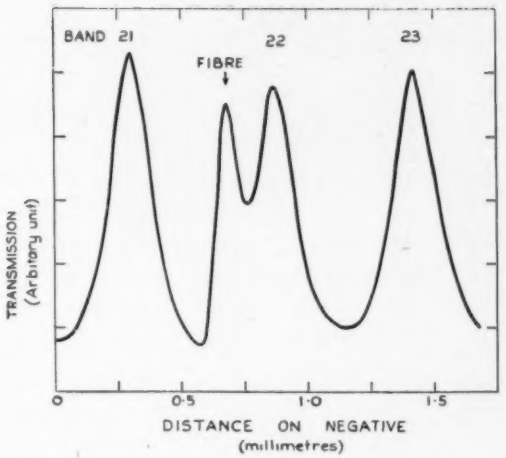


FIG. 2. Microphotometer trace across negative of Fig. 1.

We have found it possible to determine the diameters of 5 micron fibers as short as 3 mm to an accuracy better than 0.3 percent by the use of an interferometric method. Once such a measured fiber is mounted on a specimen screen it is available as a permanent direct calibration specimen suitable for magnifications up to 10,000 diameters. Magnifications up to the maximum value of 20,000 diameters of commercial instruments can be determined by the use of a 1–2-micron fiber or glass sphere calibrated by direct comparison with the interferometrically measured fiber. This small ratio of 2 to 3 may be readily and accurately determined provided allowance is made for image distortion.

It has been stated[7] that sufficient charge may accumulate on the silica fibers to cause variations in the image size when the objective current is varied during focusing. This has not been our experience, possibly because our micrographs have been taken with the condenser lens current

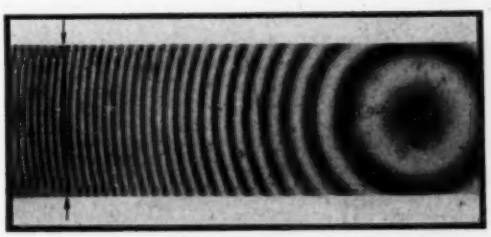


FIG. 3. Fiber No. 10 between a convex lens and an optical flat illuminated with the Hg λ5461 line.

[8] R. D. Heidenreich and V. G. Peck, J. App. Phys. **14**, 23 (1943).
[9] R. C. Williams and R. W. G. Wyckoff, J. App. Phys. **17**, 23 (1946).

TABLE I.

Method	Fiber No.	Illumination	Diameter μ	Image width (uncorrected) mm	Image width (corrected) mm	Magnification
Cover glass and flat	1A*	Na *D*	6.77	44.4	41.0	6060
Cover glass and flat	2A	Na *D*	6.18	39.3	36.8	5960
Cover glass and flat	3A	Na *D*	5.89	38.0	35.7	6060
Cover glass and flat	5A	Na *D*	5.96	38.4	36.1	6060
Cover glass and flat	9B	Na *D*	4.54	28.1	27.0	5950
Cover glass and flat	10B	Na *D*	5.91	37.4	35.2	5960
Cover glass and flat	10B	Hg λ5461	5.91			
Convex lens and flat	10B	Na *D*	5.90	37.4	35.2	5980
Convex lens and flat	10B	Hg λ5461	5.88			

* An interval of two months elapsed between the measurements of series *A* and series *B*. Visual estimates only of the band order were made in series *A*.

adjusted so that a considerable area of the wire mesh supporting the fiber is bombarded by the electron beam. This is a well-known method[10] of avoiding charging of the specimen.

I. THE MEASUREMENT OF FIBERS

The procedure in the methods we have used to measure fiber diameters involves placing the fiber between two flat glass surfaces so that it determines their separation at the fiber, and then measuring this separation interferometrically. The system is illuminated by well collimated monochromatic light, and the interference fringes are viewed by reflection with a low power microscope. The diameter d of the fiber is given by the formula $d=n(\lambda/2)$, where n is the interference band order at the fiber and λ is the light wavelength. The band order at the fiber is obtained by counting the bands back to the region of zero order where the two glass surfaces are in contact. Fractions of a band order are determined by use of a micrometer eyepiece or by photography and subsequent measurement either directly or from a microphotometer trace.

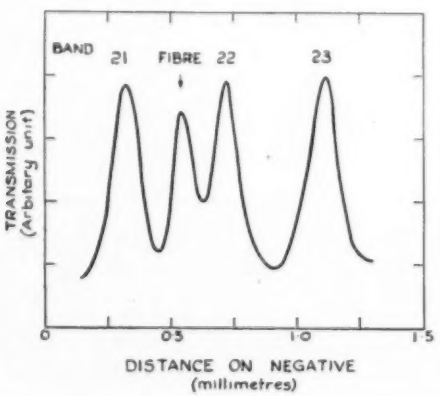


FIG. 4. Microphotometer trace across negative of Fig. 3.

In practice it is found difficult to locate a region of zero band order when thick optical flats are used, as the areas in contact are extremely small. This difficulty is easily overcome by replacing the upper optical flat with a thin microscope cover glass. Slight pressure applied a few millimeters to one side of the fiber brings the glass surfaces into contact over a considerable area. Figure 1 is a low power photomicrograph of fiber 10 (Table I) and the associated fringes obtained when such a system was illuminated by means of the Hg λ5461 line. The negative was a Kodak Super XX plate; the print was made on normal contrast paper and normal contrast development conditions were maintained. A microphotometer trace taken across the fiber shown in the negative of Fig. 1 is reproduced in Fig. 2. This curve yields 21.66 as the interference band order at the fiber which gives 5.91 microns as the fiber diameter. When the same system was similarly examined by means of the Na *D* line, the interference band order was 20.06, also corresponding to a fiber diameter of 5.91 microns.

An alternative method is to use a plano-convex lens in place of the cover glass mentioned above. The fiber is placed upon a glass optical flat with the lens adjusted so that its spherical surface touches both the flat and the fiber, the fringes are photographed, and a microphotometer trace of the negative is obtained. Figure 3 shows the Hg λ5461 fringes obtained in this manner with fiber 10. From the corresponding microphotometer trace shown in Fig. 4, the interference band order at the fiber is given as 21.54 and the fiber diameter as 5.88 microns. Use of the Na *D* line gave a band order of 20.03 and 5.90 microns as the fiber diameter. These results show that it is possible to determine the fiber position to within 0.1 of a fringe. When the Hg λ5461 line is employed the band order at a 5.5 micron fiber is about 20, so accuracy of 0.5 percent is obtainable. This is true only if the presence of dust on the

[10] J. Hillier and E. G. Ramberg, J. App. Phys. **18**, 48 (1947).

fiber and glass surfaces is scrupulously avoided. Because of the curvature of the cover glass or lens, the interference bands are not equidistant, but conditions may be readily arranged to ensure that the curvature at the fiber is insufficient to introduce an appreciable error if the fraction of the band order is determined by direct interpolation.

Tolansky[11] has found that the sharpness of interference fringes can be greatly increased by evaporating about 500A thickness of silver onto the glass surfaces. Multiple beam interference is then obtained owing to the high reflectivity of the films, and the resultant sharpening of the fringes is sufficient to allow the fiber to be located to within 0.02 of a band. Our attempts to make use of this fact were eventually abandoned for lack of sufficiently accurate means of determining the phase change occurring when light is reflected from silver films of this nature. To avoid the necessity of an accurate knowledge of the phase change, another method was tried. The fiber was held in contact with an optical flat by means of a collodion film picked up from a water surface, a silver film was evaporated onto the collodion, and observations were then made on the interference fringe system formed on illuminating the wedge between this surface and another silvered flat. The maximum displacement of the fringes as they crossed the fiber should then have given the fiber diameter to a high degree of accuracy. Attempts to observe the fringes as they crossed the fiber were unsuccessful owing to the small size of the fiber and the low light intensities available from monochromatic sources.

After measurement, the application of a little collodion cement to each end of the fiber serves to hold it permanently across a standard wire mesh specimen screen.

II. CORRECTION OF FIBER IMAGES FOR DISTORTION

As shown in Fig. 5, the images of the fibers are extremely sharp but are considerably distorted. This is largely due to the aberrations of the projection lens; the matter has been discussed at length by Hillier.[12]

[11] S. Tolansky, Proc. Roy. Soc. **A184**, 41 (1945).
[12] J. Hillier, J. App. Phys. **17**, 411 (1946).

FIG. 5. Electron micrograph of Fiber 9 showing sharp but distorted image.

One effect of this distortion is, in the case of a small sphere measured radially, to increase the magnification by a factor $f(r)$ as it is moved outwards from the image center by an image distance r. Plots (Fig. 6) of $f(r)$ against r are readily obtained from a series of exposures, taken on one plate, of a glass sphere as it is moved progressively from the field center to one corner. In order to use such a curve to correct the relatively wide fiber images some convenient method of carrying out the integration involved is required. This is apparent from the fact that an image radial distance r measured from the image center would be represented by the distance

$$r' = \int_0^r \frac{dr}{f(r)}$$

in the absence of distortion. The function $1/f(r)$ is well represented by the expression

$$1 - Ar - Br^2,$$

and the constants A and B may be determined by curve fitting to the plot of $1/f(r)$. In this way r'

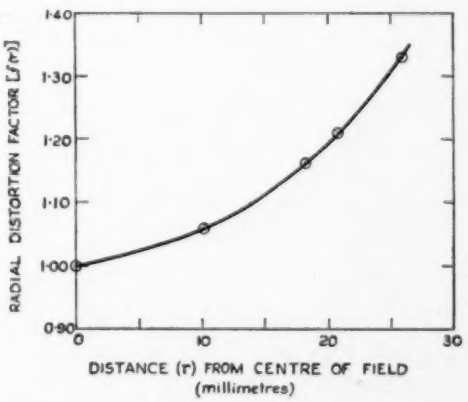


FIG. 6. Radial distortion factor $f(r)$ plotted against image distance r from center of field.

is readily calculated from the equation

$$r' = r(1 - \tfrac{1}{2}Ar - \tfrac{1}{3}Br^2).$$

The importance of this correction is shown by the fact that at a magnification of 6000 diameters the correction to a 5.9-micron fiber amounts to 6.0 percent in the case of our RCA Model U electron microscope. This instrument has been fitted with the high intensity electron gun recommended by Hillier,[13] and caution is necessary to avoid focusing too heavy a beam current onto the fibers and so fusing them. Glass spheres for the determination of distortion are readily made by feeding a coal gas flame with air or oxygen containing suspended powdered glass.

III. ACCURACY OF THE METHOD

In Table I are set out the final values for the magnification for one setting of the microscope obtained by means of six different fibers measured by the two methods described. Fibers 1, 2, 3, 5, 9, and 10 were measured between a cover slip and an optical flat, while fiber 10 was also measured between a convex lens and a flat. The deviation from the mean magnification is less than 1 percent.

The good agreement of the various results shows that the fibers are uniform and cylindrical to a high degree and that neither they nor the glass flats are seriously distorted by the pressure applied during measurement provided reasonable care is taken. It is important to realize that good reproducibility of the magnification is obtainable only if the battery potential, which controls the electron accelerating potential, is maintained constant, if hysteresis effects in the lenses are avoided by switching the lens currents off and on several times before making an exposure, and if the specimen is always supported at the same height.

IV. ACKNOWLEDGMENTS

This work was carried out at the Division of Industrial Chemistry as part of the program of the Council for Scientific and Industrial Research.

[13] J. Hillier, J. App. Phys. **17**, 307 (1946).

Electronic Radiography and Microradiography*

JEAN-JACQUES TRILLAT
Directeur du Laboratoire de Rayons X du Centre National de la Recherche Scientifique, Paris (Bellevue), France
(Received February 27, 1948)

Two techniques of radiography by means of secondary photoelectrons liberated by hard x-rays (180 kv), one by reflection and one by transmission, are described. Special fine-grained photographic emulsions of the Lippmann type are required which will not be affected by x-rays traversing the film, but are sensitive to photoelectrons liberated from the surface of a specimen in intimate contact with the photographic film or paper. In this way photographic negatives may be reproduced (from the silver of the image), microradiographs prepared by enlargement of any area of the radiograph, surface topography studied, ores analyzed, etc. For electronic radiography by transmission, instead of the specimen being its own electron generator, a thin metal foil (usually lead) pressed against the specimen generates the electrons which then traverse the thin specimens, and register on the photographic film or paper. Applications are papers, insect wings, varnish or rubber films, plant leaves, tissue sections, plastics, and alloys.

THE two methods described herewith are totally different from the usual radiographic processes because they make use of secondary photoelectrons instead of x-rays as a means of registration of images on the photographic emulsion. We shall call these methods electronic radiography by reflection and electronic radiography by transmission, respectively.

* Edited, submitted for publication, and proofread by G. L. Clark, University of Illinois, Urbana, Illinois.

I. ELECTRONIC RADIOGRAPHY BY REFLECTION

Principle

In April 1940, in a paper published in the Revue Scientifique,[1] it was suggested that a method designed to obtain radiographs or microradiographs from finished articles without having recourse to a destructive method of sample preparation should prove useful and interesting.

The method developed to achieve this purpose, designated "method by reflection," is based on the following principle.

Let us consider a plane material surface (having any dimension) of an article in which components have quite different atomic numbers. When irradiated by x-rays having adequate wave-lengths, the surface becomes the source of a secondary scattered fluorescent wave emission and simultaneously of a photoelectron emission whose kinetic energy, according to Einstein's and M. de Broglie's equation, depends upon the incident quantum and the extraction work of the electrons, that is to say, upon the atomic number:

$$\tfrac{1}{2}mv^2 = h\nu_{(\text{incident})} - W_n,$$

where W_n is the extraction work corresponding to an n level; it varies proportionately with the atomic number.

The number of emitted photoelectrons depends upon the wave-length and the intensity of the incident x-rays. It increases when the atomic number increases, heavy atoms actually having more electrons than light ones. Thus, use is made of the photographic effect of the secondary electronic emission which is closely related to the fluorescent phenomenon. It should be noticed that this is merely a surface effect, electrons being much more absorbed than x-rays in a material substance.[2–5]

In order to obtain the desired result, it is necessary that the emulsion which is in closest

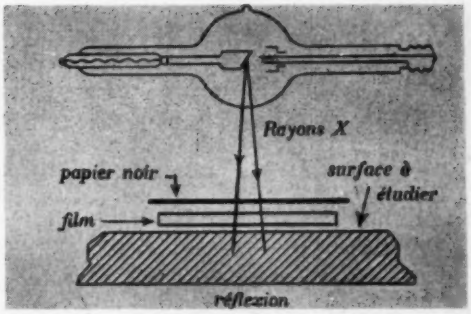


FIG. 1. Diagram of the method of electronic radiography by reflection.

possible contact with the surface to be investigated and through which the incident x-rays will pass *first* should be back impressed by the photoelectrons only and not by the incident or simply scattered x-rays. It is necessary that the absorption coefficient of the photographic emulsion should be very low for this particular radiation and, on the other hand, that the photographic emulsion should be impressed by the photoelectrons in a sufficiently strong way.

In order that the absorption coefficient of the emulsion for x-rays should be low, it is necessary that the photographic layer be very thin; it must consist in very fine and non-sensitive grains. Furthermore, the x-rays used must be hard, that is to say, they must be emitted by a tube operating at high voltage; thus the absorption in the emulsion and therefore the photographic effect are lessened.

The first condition is realized when one uses special Lippmann films which are used in microradiography, the thickness of the sensitive layer being approximately 1/300 mm. Excellent results have been obtained with Lippmann films prepared by Eastman Kodak Company. Thus, highly contrasting emulsions may be employed and even silver chloride photographic papers

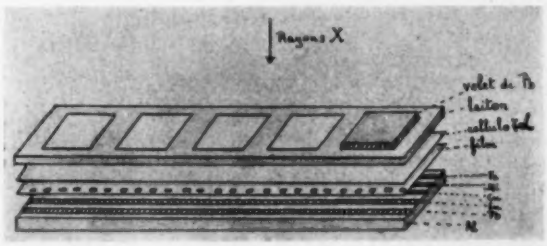


FIG. 2.

[1] J. J. Trillat, "La microradiographie," Rev. Sci. **4**, 212 (1946).

[2] J. J. Trillat, "Sur l'action photographique des electrons secondaires resultant de l'action des rayons x sur les metaux," Comptes Rendus **216**, 179 (1943).

[4] J. J. Trillat, "Sur une extension de la methode de radiographie par reflexion," Comptes Rendus **216**, 683 (1943).

[5] J. J. Trillat, "Radiographie et Microradiographie par transmission et par reflexion," Bin Ste Fr. Electriciens **6**, III, 1–20 (1943); and Rev. Sci. **8**, 364, 370 (Sept., Oct. 1942).

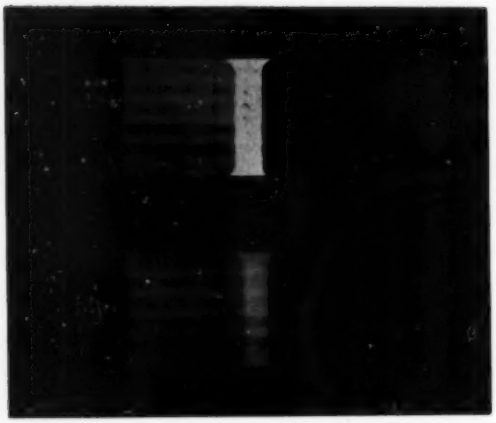


FIG. 3.

(Velox) which are only slightly sensitive to x-rays but very sensitive to electrons.

The second condition is satisfied when using constant voltages higher than 100 kv and up to 200 kv.

It may be seen that this method is quite different from ordinary microradiography or radiography by transmission, as illustrated by Fig. 1, in the following respects:

(a) The photographic film is first traversed by incident x-rays (which looks like a paradox).

(b) X-rays are not active, but photoelectrons are.

(c) Instead of using soft x-rays as is generally the case with ordinary microradiographs, hard x-rays are used here.

(d) Instead of investigating the whole of a radiated mass which results in the appearance of superposed layers, the structure of a purely superficial film is investigated.

The technical procedure is very simple; it comprises applying the film very closely against the surface to be investigated and exposing the whole system to x-rays. It is essential that the film should be in very close contact with the surface because photoelectrons are emitted in all directions and are greatly scattered in air.

Results

(a) *Investigation of metallic surfaces*

Let us place in an adequate holder (Fig. 2) a Lippmann film in contact with a plane surface consisting of metals having different atomic numbers (26 Fe—28 Ni—29 Cu—50 Sn—82 Pb) and let us expose the surface to x-rays which are emitted by a tube operating under 180 kv; the soft part of the continuous spectrum is eliminated by means of a 4-mm copper filter. Under these conditions, a picture is obtained (Fig. 3) where the photographic density measured with a Capstaff-Purdey densitometer is exactly proportional to the atomic number as shown in Table I.

Thus, elements of consecutive atomic number such as Fe, Ni, and Cu, are separated easily and one may foresee interesting applications of this fact in the field of alloys and mineral studies.

Provided a certain care is taken, it is even possible to use ordinary commercial fine-grained

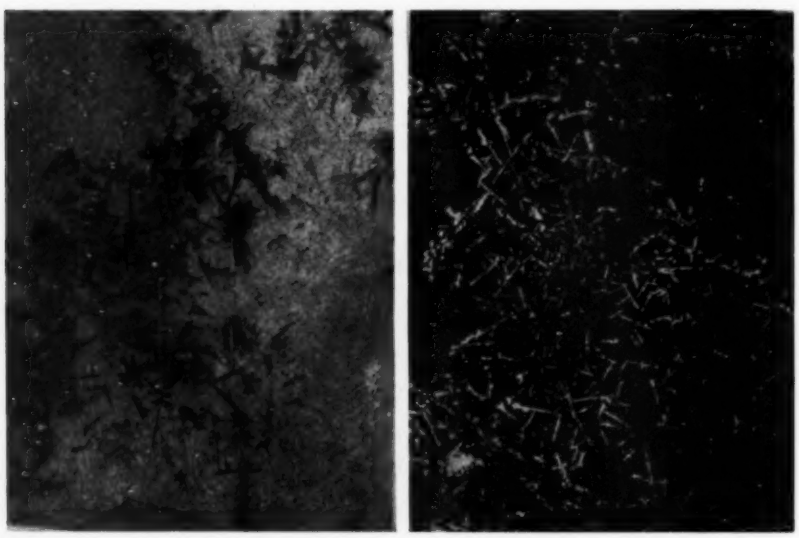


FIG. 4. Comparison of an ordinary radiograph of an alloy of aluminum, antimony, tin, and lead and an electronic radiograph by reflection of the same alloy (Saulnier). The ordinary radiograph is made at 50 kv, the electronic at 180 kv; filtered 2-mm complex and 2-mm aluminum; exposure time 4 min.; distance 45 cm.

FIG. 5. Uranium mineral, enlargement 6 diameters.

FIG. 6. Pitchblende, black uranium, gray lead, white SiO_2, enlargement 6 diameters.

emulsions, such as "Panatomic" Kodak emulsion. Exposure times are greatly reduced and the process requires a few seconds only; when such films are adopted, the field of application of the method is extended; true radiographs by reflection are made and then magnified 20 or 30 times, later on.

Systematic tests have shown that the optimum practical range of voltages for this method is 150–200 kv; in this range the effects of fluorescent x-rays are largely negligible in comparison with the photoelectron effect. Ordinary microradiography by transmission would not have given the same results without using monochromatic x-rays and selecting the radiations so that their wave-lengths should fit inside the critical absorption discontinuities of the elements to be differentiated.

(b) *Alloy inspection*

When using Lippmann films, one may convert the radiographs by reflection into as many microradiographs as required; in fact, each point of the film might be magnified many thousands of times because of the excessive fineness of the grain; thus beside the whole image, any required detail photographs are obtained. Thus, a single process is sufficient to yield simultaneous information about the macro- and the microstructure of a specimen.

TABLE I.

Fe	$N=26$	Photographic density 0.83
Ni	28	0.85
Cu	29	0.86
Sn	50	0.94
Pb	82	1.10

Figure 4 illustrates a radiograph by reflection of an Al alloy containing Pb, Sb, and Sn.

In case macrographic investigation only is aimed at (for instance, for segregation or heterogeneity investigations), it is not necessary to use a special very fine-grained film, since the image obtained requires no magnification. Excellent results were obtained by simply placing a relatively insensitive photographic silver chloride paper, for instance a "Velox contrast" paper, in contact with the surface. All manipulations may then be made without resorting to complete darkness, and exposure time is greatly reduced.

(c) *Reproduction of photographic negative*

The silver image of a film or plate negative might be used as a secondary electron source

FIG. 7. Galena containing nickel arsenide and SiO_2, enlargement 6 diameters.

which will register an image on a fresh photographic paper of film.[2] A negative to be reproduced is placed at the bottom of an ordinary photographic frame in which the glass has been replaced by a black celluloid sheet; next to the negative is placed a fresh photographic paper or film with its sensitive surface in contact with the silver image. The whole system is exposed to hard x-rays which first go through the celluloid sheet, then through the paper, and finally strike the negative plate. Photoelectrons emitted by the silver of the image will then impress the photographic paper which on development will yield the *exact* image of the negative. This reproduction process gives a direct copy of the original without any negative intermediate.

(d) *Surface topography investigation*

The reflection method enables us to get some information concerning the mechanical condition of a surface. When a surface has been mechanically worked and when superficial heterogeneities exist, these defects will appear prominently on the exposed negative. The process is reliable only when the path in air is not too long, on account of electron scattering.

(e) *Study of ores*

Lastly, we have recently applied this method to the study of polished surfaces of various ores containing elements with different atomic numbers. The results were excellent, and it is to be supposed that this process will be of use to geologists and mineralogists.[6]

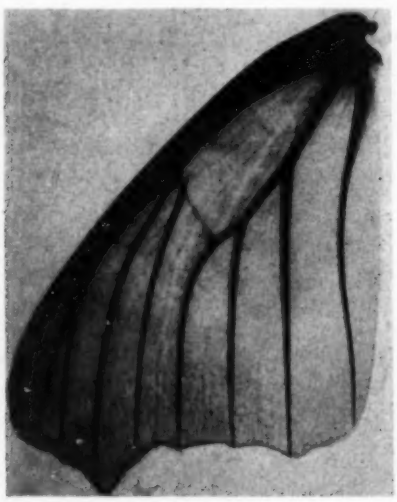


FIG. 8. Electronic radiograph by transmission. Butterfly wing, enlargement 8 diameters.

For example, Figs. 5 and 6 represent a piece of uranium ore (pitchblend) in which the uranium present is clearly outlined in black. The bright parts correspond to the lighter elements such as silicon.

Figure 7 represents ore containing plumbago (PbS), together with nickel arsenide ($NiAs_2$). The whites correspond to the gangue of silica (SiO_2), the blacks to the plumbago, the grays to $NiAs_2$. This method can even be used to find out the quality of the ore under examination, by measuring the blackening of the plate and comparing with a standard one.

Limitations of the Electronic Radiography by the Reflection Method

Limitations of this method depend on the electron diffusion (scattering) almost solely, since electrons are emitted in any direction by the surface layer; however, the inconvenience is not so great as when one uses the fluorescent secondary x-rays, as M. Guinier did.[7] Nevertheless, it is possible to minimize the scattering nuisance by applying the photographic film or paper as perfectly as possible against the investigated surface and by using a vacuum apparatus when necessary. In any case, it is absolutely essential to have a nearly perfect contact between the object and the film, because of the great absorption and diffusion of electrons in air.

Discussion of Results

It was explained above that secondary electrons are active in our experiments and fluorescent x-rays are not. Scattered x-rays, both unmodified and modified by the Compton effect, do not appreciably affect the film. On the contrary, secondary fluorescent x-rays which are emitted upon impact with primary quanta—K, L, etc., rays—may possibly have fairly long wave-lengths, especially when light elements are involved, and thus might well play a part in the

[6] J. J. Trillat and Ch. Legrand, "Radiographie electronique de minerals," Comptes Rendus **224**, 1000 (1947).
[7] A. Guinier and J. Devaux, "Sur la radiographie par reflexion," Comptes Rendus **214**, 223 (1942).

investigated effect on account of their rather high absorption.

Within the experimental conditions, characterized by voltages higher than 100 kv, we may safely conclude that the film is affected by photoelectrons only. This may be checked by a very simple experiment. With a surface consisting of aluminum, iron, or lead, an exposure is made first by reflection of electrons generated below 180 kv; then a second exposure is made with conditions being exactly the same except that a very thin celluloid foil (1/20 mm) is placed in the path of the x-ray beam. In the last case, no image will appear, a proof that we are actually concerned only with an electronic effect. If characteristic x-rays, such as copper, iron, or lead K-rays which are strongly excited in these conditions, were involved in the effect, the low thickness of the celluloid foil (1/20 mm) would not stop them, and a blackening of the film almost as severe as the one observed in the first experiment would be observed.

It may be interesting to compare these results with those obtained later by M. Guinier,[8] who made use of long wave-length incident x-rays for his researches instead of using short wave-length x-rays, as was done in this work.

M. Guinier has shown that when illuminating an iron plate with rays having a wave-length below that of the K absorption discontinuity of iron ($\lambda_K = 1.74$A), an intense blacking will appear on the photograph; it is due to fluorescent K-rays of iron. But when wave-lengths are shorter, for instance, about 0.25A (that is to say, the voltage is 50,000 v) there is a change in the phenomenon and the effect is mainly due to photoelectrons; this is checked by interposing a 2/100-mm thick Cellophane foil in the path of the rays. Our own experiments which were carried on partly at 70,000 to 100,000 v demonstrated that photoelectrons only were decidedly active.

Thus, the observed effect represents the combination of two other effects, the relative magnitude of which depends on the incident radiation wave-length; when taking into account the number of photoelectrons, their energy, and decreasing sensitivity of the photographic emulsion as wave-length decreases, it will be found

[8] A Guinier, "Etude theorique de la radiographie par reflexion," Revue Sci. **63**, 52 (1943).

FIG. 9. Electronic radiograph by transmission of a sheet of cigarette paper, enlargement 8 diameters.

easily that the combined effect is due solely to fluorescent x-radiation when $\lambda < \lambda_K$. This effect remains practically constant for medium wave-lengths, and then increases when wave-lengths become shorter. Then, electrons only act on the emulsion; finally, the effect decreases when rays of high penetrative power (above 200 kv) are used. M. Guinier's experiments constitute a favorable complement to our own, in the long wave-length field. Analytical applications may be foreseen for the method of using the fluorescence radiation of illuminated elements instead of photoelectrons.

II. ELECTRONIC RADIOGRAPHY BY TRANSMISSION

Working with M. A. Saulnier,[9] the author has also attempted to use the photoelectric effect in order to obtain radiographs from a very thin

FIG. 10. Electronic radiograph of a section of rabbit kidney.

[9] A. Saulnier and J. J. Trillat, "Radiographie par electrons secondaires," Comptes Rendus **220**, 772 (1945); and Revue Sci. **3244**, 211 (1945).

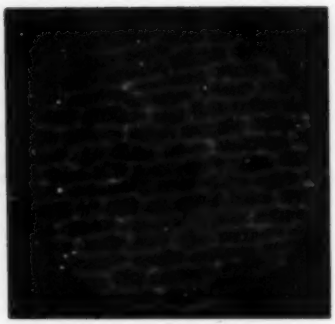


FIG. 11. Electronic radiography by transmission of an onion peeling.

object; instead of being its own electron generator, as in the radiography by reflection, the article is traversed by photoelectrons, extracted by means of x-rays from a metallic surface against which the article is strongly pressed.

To obtain the required result, it is necessary that the photoelectrons should be emitted in great quantity and that they should have sufficient energy to go through a preparation which may be a few thousandths of a millimeter thick; these conditions are practically obtained when a lead radiator is used and when a 150- to 200-kv primary quantum is used. It is, of course, necessary for the thin sheet or film to have a constant thickness, any local thickness variation being made obvious by an additional absorption of the photoelectrons.

Incident x-rays produced in a tube at 150–200 kv first go through a blackened celluloid plate which acts to prevent light from striking the film, and for maintaining the investigated preparation. Then, radiation is passed through a perfectly smooth and plane 0.2-mm thick lead foil which filters out the soft part of the rays and which emits from its back surface a large quantity of photoelectrons, thus playing the part of a radiation-transmitting unit. The preparation to be inspected is tightly held between the radiating source (lead foil) and a photographic film or paper. Under these conditions, the photographic emulsion is affected by the photoelectrons which passed through the preparation and also by those which are emitted by the preparation proper. Incident or fluorescent x-rays do not produce any noticeable image, as shown by our previous researches, provided a tube operating between 150 and 200 kv is used.

FIG. 12. Bubbles of air in a film of paint of different thickness. The bubbles appear as the black dots, enlargement 3.5 diameters.

In this way a true electron radiography is obtained, and it may be compared with the results given by an electron microscope, if magnifying power is overlooked. When an extremely fine-grained emulsion is used, such as the one used in microradiography, the resulting image might be magnified and many details which did not appear on the original print are made clear. Exposure time varies from a few seconds to several minutes.

This method has been applied successfully on papers, insect wings, varnish or rubber films, plant leaves, plant or animal tissue sections, plastics, alloys, etc. Very clear pictures are often obtained and give added information to the microscopic inspection; this result is interesting in particular when opaque specimens which cannot be inspected under a microscope are concerned. A few examples are given in Figs. 8–11. I have also applied this method to the study of thin oil films which are in contact with pieces of machinery. It is thus possible to disclose the structure of films about 2μ thick. Further study of some lubrication phenomena and paints[10] (Fig. 12) is in progress. It is quite obvious that the examples could be multiplied and the method applied to numerous cases. It is a very simple method which requires only ordinary medical radiographic apparatus, provided it may be operated above 150 kv. The x-ray apparatus need not even be kept at a constant voltage; a pulsating voltage is sufficient provided the radiation is filtered through a few millimeters

[10] J. J. Trillat and Ch. Legrand, "Sur une application de la Radiographie electronique a la detection de couches minces organiques ou minerales," Comptes Rendus **224**, 645 (1947); Rev. de l'Institut Français du Petrole **2**, 250 (1947).

copper so as to eliminate the soft part of the spectrum.

We must point out that a quite similar method was lately utilized in Great Britain by H. S. Tasker and S. W. Towers;[11] Our own researches had been initiated for a long time when we heard of the work which was performed quite independently. The English authors obtained results which quite agree with our own.

Discussion of Results

Electronic radiography by transmission is based on the absorption of highly energetic photoelectrons by matter; it is thus essentially different from soft x-ray radiography such as has been utilized by Dauvillier[12] and Lamarque.[13] The photoelectron absorption law is very different from that of x-ray absorption and is less well known; while x-ray absorption varies proportionally to the fourth power of the atomic number, the electron permeability of matter which depends on the density has no simple relation to the atomic number. It varies little from one element to another when compared with the coefficient of absorption of x-rays; thus electron absorption in gold is nearly three times that in aluminum, while the x-ray absorption coefficient is twenty times greater in gold. "Electronic" negatives should then be explained in a different way than x-ray negatives; exactly the same problem arises in the explanation of images obtained with the electron microscope by means of electrons having essentially the same energy as that of the photoelectrons in our researches which pass through a very thin specimen.

As might be expected, the main limitation to the method is the thickness of the specimen; it must be very thin (a few to several thousandths of a mm) and the thickness must be constant. A. Saulnier[14] made measurements to show that the absorption of photoelectrons by aluminum or Cellophane screens does not depend on the quality of the absorbant matter, but in a rough approximation on its density only.

Moreover, photoelectrons are emitted in any direction and with varying energy from the lead surface; thus, the sharpness of the image might be altered.

In order to determine the best conditions for applying the method, it is, of course, necessary to investigate theoretically the secondary electron emission conditions. This research was executed by M. A. Saulnier and will be the subject of a publication elsewhere. It was found, for instance, that electrons ejected backwards (Compton effect) need not be taken into account, that photoelectron emission is about 1.2 times stronger along the direction of the incident radiation direction than in the opposite direction (Cooksey), and that the asymmetry increases as wave-length decreases and atomic weight decreases. The computation shows that lead surface layers only ($\sim$8/100 mm for L electrons) are involved in the emission.

Finally, it is interesting to note that electronic radiography is produced from hard x-rays which are entirely unsuited for the radiography of such thin articles directly, but that the lead screen, acting by converting the hard rays into corpuscular rays which are strongly absorbable and possess a very high photographic effect, makes possible a highly successful technique.

SUMMARY

When x-rays hit an element of atomic number N, it is a known fact that the kinetic energy of the photoelectrons produced depends on the incident quantum (by application of Einstein's law).

These photoelectrons are strongly absorbed by the photographic emulsion which they deeply affect when their energy is sufficient.

Under these conditions, in order to study any surface made up of elements having different atomic numbers (an alloy, for instance), it is sufficient to place a film or photographic paper in contact with this surface and to irradiate the whole with a beam of hard x-rays (150 to 200 kilovolts) which affect only very slightly the silver bromide because of their weak absorption. An image of this surface is thus given by electrons, a picture in which the different elements come out in different shades of black; superficial irregularities can also be thus revealed. By

[11] H. S. Tasker and S. W. Towers, Nature **156**, 50 (1945).
[12] A. Dauvillier, Comptes Rendus **191**, 1287 (1930); J. Phys. **6**, 229 (1932).
[13] P. Lamarque, "L'Historadiographie, Union Intern. contre le Cancer," **4**, 205 (1939).
[14] A. Saulnier, "Adsorption des electrons secondaires par des ecrans minces," Comptes Rendus **220**, 772 (1945).

making use of grainless films, of the Kodak-Lippmann type, for instance, each point of the photograph can be enlarged and details obtained.

On the other hand, the photoelectrical emission from high atomic number metals, such as lead, can also be used for the purpose of examining the internal structure of thin films such as paper, vegetable or animal tissues, paintings, etc. The object under study will then be placed between a sheet of lead acting as radiator and a photographic film, and the whole is submitted to the action of hard x-rays; the photoelectrons radiated by the lead register on the film after passing through the object. This last process has been used jointly with Dr. Saulnier.

Examples of applications are given by the author together with suggestions of further development of these methods.

A Temperature-Controlled Frevel Focusing X-Ray Powder Camera

V. VAND
Research Department, Lever Brothers & Unilever, Limited, Port Sunlight, Cheshire, England
(Received February 20, 1948)

A focusing x-ray powder camera of Frevel type is described, suitable for the study of large cells of organic crystals. It uses a convergent x-ray beam collimated by a multiple slit system and records transmitted reflections only.

INTRODUCTION

A GREAT shortening of the time of exposure can be obtained by the use of "focusing" methods in x-ray diffraction. The Seeman[1]-Bohlin[2] arrangement is well known for back-reflection diffraction patterns. Frevel[3] uses a convergent Soller slit for focusing the x-ray beam, the geometry of the arrangement being otherwise that of Guinier[4] and Johansson,[5] which uses a bent and ground quartz crystal for focusing. Other arrangements are those of Smith.[6]

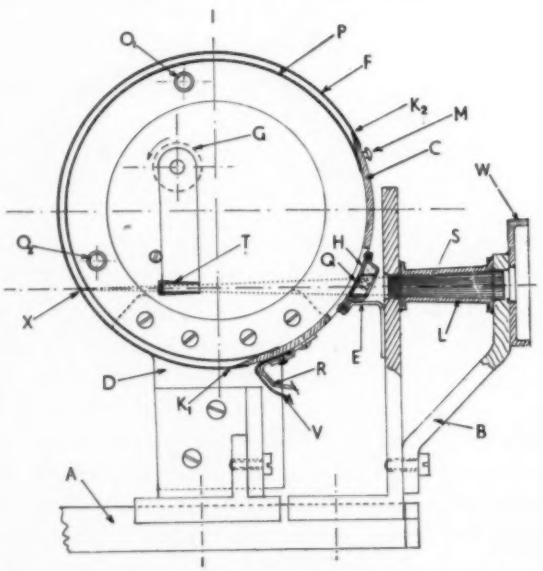


FIG. 1. Vertical cross section of the 12.5-cm diameter focusing x-ray powder camera.

In the Frevel arrangement, the diffraction pattern is produced on transmission through a thin sheet of the specimen, and it is thus suitable for the study of organic compounds, having large unit cells.

In order to fulfill the focusing condition, the film, the specimen sheet, and the focus must all lie on a common circle.

GENERAL FEATURES OF THE 12.5-CM FOCUSING CAMERA

In this paper a focusing camera of Frevel type is described which uses a convergent beam collimated by Soller slits, arranged to converge to a linear focus. The advantage of the Guinier's camera of the monochromatic radiation is lost, but the camera is faster and can easily be made from available materials, which compensates this disadvantage. The camera and its components are shown in Figs. 1, 2, and 3.

The camera consists of an adjustable base *A*,

[1] H. Seeman, Ann. d. Physik **59**, 455 (1919).
[2] H. Bohlin, Ann. d. Physik **61**, 421 (1920).
[3] L. K. Frevel, Rev. Sci. Inst. **8**, 475 (1937).
[4] A. Guinier, Comptes Rendus **204**, 115 (1937); J. Sci. Inst. **22**, 139 (1945).
[5] T. Johansson, Zeits. f. Physik **82**, 507 (1933).
[6] C. S. Smith, Rev. Sci. Inst. **12**, 312 (1941).

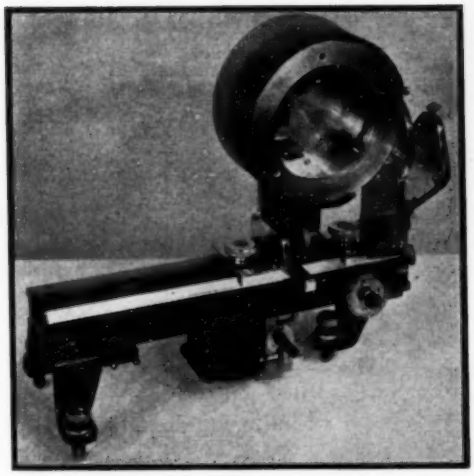


FIG. 2. The camera assembled.

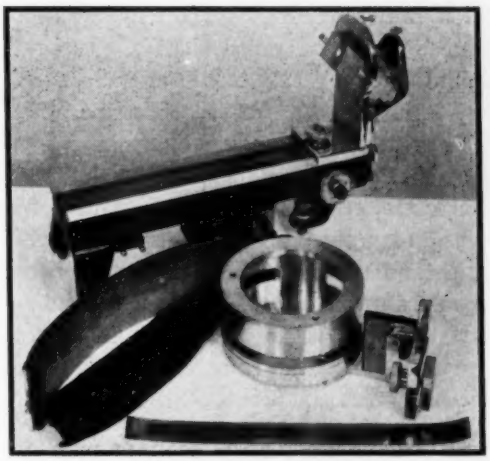


FIG. 3. The camera opened.

holding a bracket B, with a multiple slit system S and a temperature controlled curved sheet specimen holder H carrying the specimen Q, a bracket D with a hollow cylinder C holding the film F in position and a beam trap T. The base is provided with three leveling screws; for this purpose, we have used the base of a Metrovick Universal Camera.

THE COLLIMATING SYSTEM

The collimating system S consists of twenty slits, made to converge to a focus X at an approximate distance of 12.5 cm from their exit ends. The slits are constructed from a pack of twenty-one strips, spaced from each other by pairs of wedge-shaped spacers. The strips used were 4.3 cm long, 1 cm wide, and 0.1 mm thick. The spacers were each 4.3 cm long and 2 mm wide, placed at each side of the strip, which leaves 6 mm for the height of each slit. The thickness of the spacers was 0.075 mm at the exit end and 0.135 mm at the entrance end, which gives the correct angle of the wedge to produce a focus at 12.5 cm from the exit of the slit system. The strips were made from razor blades, which are sufficiently uniform and flat, and the spacers by folding a sheet of aluminum foil 0.025 mm thick in a steplike manner. The foils may be cemented together with a trace of lacquer and strips 2 mm wide may then be cut from the resulting wedge. During the assembly, the wedges and the steel strips may be held in position by applying a trace of lacquer to their surfaces. After the assembly, the whole pack is enclosed in a lead shield L. The half-width of the beam at the focus X obtained by the above arrangement is about 0.4 mm.

As a precaution against fouling the ends of the slits S by the beams, the exit ends of the steel strips were made slightly thinner by grinding, serving thus as the side pieces of the ordinary camera slits.

THE SPECIMEN MOUNTING

The specimen Q is placed in a curved specimen holder H, which in turn is located by two pins on a temperature-controlled bracket E. The bracket is hollow and is maintained at any temperature between 0° and 100°C by water circulated from a thermostatically controlled bath by a pump. The temperature of the specimen is not exactly that of the water, so that a calibration curve must be prepared; melting points of pure fatty acids can conveniently serve as calibration temperatures in the range used.

The specimen holder H is made of a sheet of metal 1 mm thick, which is bent to a radius 6.25 cm and provided with a hole of the same diameter as the incoming x-ray beam. If the specimen Q is self-supporting, it is fixed or pressed into the hole; otherwise the specimen may be mounted between two thin Cellophane sheets, fixed at each side of the holder. It is important to have a specimen of correct thickness. A thickness equal to or thinner than $1/\mu$, where μ is the linear absorption coefficient, should be used. For

heavier minerals, etc., specimens of a correct thickness may be conveniently prepared by sedimentation of a suspension of the specimen in methyl ether with a drop of collodion added, in a Petri dish, at the bottom of which was previously placed a Cellophane strip. After the evaporation of the ether, the strip with the sediment attached to it may be peeled off the bottom of the dish and transferred to the specimen holder. By measuring the area of the dish and by weighing the amount of the material added, the absorption can be accurately calculated from the mass absorption coefficient of the material.

THE FILM ARRANGEMENT

The film F is pressed against the outside of the cylinder C by means of a strip of velvet V, pressed on the film by a rubber band R. The rubber band is hooked on a rivet M.

A standard film 3.5 cm$\times$28 cm is used. No paper casette is used, and the light is excluded by the velvet V and by a strip of black paper P fixed inside the brass cylinder. The paper can be, however, replaced by a nickel foil to serve as a filter for CuK_α radiation.

The cylinder is provided with windows which admit the convergent beam from the outside and which expose a suitable strip of film to diffracted beams. The center-line of the beam does not need to pass through the axis of the cylinder, so that the camera can be considerably offset and the film placed asymmetrically. This extends the range of the angles registered on one side of the film up to $2\theta=90°$. The pattern produced by the arrangement is independent of the amount of the offset, so that great accuracy is maintained without any need for fine adjustment of the beam, provided the film and the specimen lie accurately on the same circle.

The cylindrical body of the camera is provided with a lid (not shown on the photographs) and with openings O_1 and O_2, through which the camera can be filled with hydrogen.

THE BEAM TRAP

A lead beam trap T is placed about 3 cm in front of the focus X. At the bottom of the trap is fixed a small fluorescent screen, which serves as a useful check of the alignment of the camera. It is not practicable to place the beam trap too close to the specimen, in order to reduce air scatter, as the broadness of the beam would require a large trap. The trap is on a hinge and it is convenient to swing it out of the beam, by means of an outside knob G, for a short time to register the central spot, which serves as a starting point for measurements of the spacings. Knife edges K_1 and K_2 are provided at either end of the window covered by film to determine film shrinkage.

PERFORMANCE OF THE CAMERA

The camera is most economically used in conjunction with a broad x-ray source, so that the whole slit system is filled with the radiation. This can be obtained by de-focusing a demountable x-ray tube. This lowers the specific loading of the target and makes it possible to run the tube at a slightly higher power.

From the geometry of the arrangement, it follows that the camera gives a pattern on exactly twice the scale of an ordinary powder camera of the same diameter.

As the reflections are integrated from a comparatively large area of the specimen, no rotation or oscillation of the specimen is necessary. Typical exposure times are 1 hour with CuK_α radiation filtered with Ni foil at 45 kv and 20 ma.

For smaller Bragg angles, the relative intensities of diffracted lines should be approximately the same as in a conventional powder camera; for large Bragg angles, there is a strengthening of line intensities due to the x-rays penetrating the film at an angle, which is, however, compensated to a certain degree by weakening of the beams due to traveling a longer path through the specimen. The relative intensities are also affected by preferred orientation in the specimen. This can, however, be detected by comparing line intensities on the right and left half of the film which are then different because the specimen is inclined to the beam. The errors in measurements of spacings arise from the following factors:

(1) film shrinkage, which can be corrected by the use of the knife edges,

(2) effective camera radius, which is slightly larger than the cylinder radius owing to the finite thickness of the film,

(3) inaccuracy of setting of the specimen on the common circle and inaccuracy of reproducibility of this setting,

(4) for thick specimens, the effective depth from which X-rays are diffracted varies very slightly with the Bragg angle and specimen thickness, and

(5) the diffracted lines are formed by a superposition of many circles, the centers of which all lie along the focal line. This tends to shift point of gravity of the diffracted lines towards smaller Bragg angles.

The errors arising from points 2, 3, and 5 can be largely eliminated by calibration. The camera was calibrated by using standard substances, such as quartz, rocksalt, sugar, and gypsum; the calibration was extended to smaller Bragg angles by using the lower orders of the long spacings of a homologous series of silver soaps, which give exceptionally sharp reflections.

The calibration proved that the errors of the camera are not serious, or can be eliminated by careful experimental technique; indeed, the camera seems to be more free of systematic errors than conventional cameras.

After application of corrections, the probable error of a single line measurement was found to be of the order 1 percent at 20A and 0.1 percent at 2A, using CuK_α radiation.

ACKNOWLEDGMENTS

The author wishes to thank Dr. J. Iball for valuable advice in constructing the camera, to Mr. T. R. Lomer for calibration of the camera, to Mr. E. T. Sanders for making the components, and to the Directors of Lever Brothers and Unilever, Limited, for permission to publish this paper.

The Absorption of Incident Quanta by Atoms as Defined by the Mass Photoelectric Absorption Coefficient and the Mass Scattering Coefficient

JOHN A. VICTOREEN

The Victoreen Instrument Company, Cleveland, Ohio

(Received May 11, 1948)

Evidence presented indicates that the absorption of quanta as defined by the mass photoelectric absorption coefficient τ/ρ and the mass scattering coefficient σ/ρ of any element may be represented by a simple expression.

Tabulated values of the calculated mass absorption coefficients μ/ρ are given for wave-lengths from 0.01 to 40 angstroms for hydrogen and carbon and from 0.01 to λ_K and λ_K to λ_L for aluminum, copper, tin, and lead. Graphs are given showing agreement with observed values.

INTRODUCTION

THE author, in a previous article,[1] presented an expression for the calculation of absorption coefficients. The data were purely empirical in that every attempt was made to represent observed values and, when necessary, to compromise differences between observers. It was believed, however, that a simple law could be derived from the empirical data. This appears to have been done satisfactorily, but as might be expected from such an approach to the problem facts may be indicated before they are fully explained. Thus, the quantum states required to express the absorption of the elements have been found from agreement with observed data.

In the present article an attempt has been made to preserve quantum conditions at the expense of the observer. It appears that this has improved the general agreement with observed data and made the individual observer's errors more evident. Examples of this may be seen in Fig. 3 for aluminum and Fig. 4 for copper. A change in the quantum numbers assigned will improve the agreement at wave-lengths approaching λ_K for copper and aluminum, but on close inspection one may possibly be justified in assuming the given calculated values to be correct.

The generally excellent data of Allen[2] are shown to contain a large wave-length error on wave-lengths below 0.2A and have not been used below this in this work. Corrected, but unpublished data presented to the author by S. J. M.

[1] J. A. Victoreen, J. App. Phys. **14**, 95 (1943).

[2] A. H. Compton and S. K. Allison, *X-Rays in Theory and Experiment* (D. Van Nostrand Company, Inc., New York, 1935), see Allen's tables.

TABLE I. Calculated mass absorption coefficients wave-length in angstroms.

λ	H 1	C 6	Al 13	Cu 29	Sn 50	Pb 82
0.010	0.1134	0.0571	0.0551	0.0523	0.0490	0.0492
0.012	0.1240	0.0624	0.0602	0.0573	0.0542	0.0563
0.015	0.1388	0.0694	0.0675	0.0644	0.0620	0.0688
0.020	0.1577	0.0794	0.0767	0.0739	0.0743	0.0944
0.025	0.1736	0.0874	0.0845	0.0826	0.0880	0.1298
0.030	0.1870	0.0941	0.0910	0.0907	0.1040	0.1778
0.040	0.2083	0.1052	0.1024	0.1073	0.1466	0.3210
0.050	0.2252	0.1135	0.1112	0.1255	0.2078	0.5403
0.060	0.2407	0.1215	0.1201	0.1486	0.2942	0.8507
0.080	0.2628	0.1331	0.1353	0.2102	0.5581	1.780
0.100	0.2794	0.1419	0.1502	0.3018	0.9733	3.163
0.12	0.2926	0.1494	0.1672	0.4323	1.571	*5.010*
0.15	0.3081	0.1592	0.1985	0.7180	2.871	2.040
0.20	0.3263	0.1741	0.2740	1.495	6.324	4.535
0.25	0.3366	0.1884	0.3882	2.741	11.49	8.459
0.30	0.3498	0.2089	0.5573	4.561	18.61	14.05
0.40	0.3599	0.2588	1.087	10.26	*38.46*	30.86
0.50	0.3660	0.3357	1.950	19.25	12.21	55.74
0.60	0.3736	0.4494	3.226	32.02	20.55	88.60
0.80	0.3854	0.8119	7.318	70.38	46.37	
1.00	0.3939	1.402	13.95	126.9	86.26	
1.20	0.4069	2.280	23.68	*201.4*	141.7	
1.50	0.4318	4.256	45.24	45.83	255.6	
2.0	0.4927	9.765	103.6	105.0	522.2	
2.5	0.5916	18.76	195.6	198.2	856.6	
3.0	0.7375	32.10	326.7	331.2		
4	1.208	74.92	720.7	730.7		
5	1.984	144.3	1304	1321		
6	3.142	246.2	*2072*	2100		
8	6.902	568.9	279.3	4112		
10	13.09	1082	536.9	6340		
12	22.32	1821	912.7			
15	43.17	3413	1739			
20	101.6	7522	3950			
25	197.6	1357	7377			
30	340.7	2154				
40	804.0	4198				

Allen are in general agreement with Cuykendall[3] and Jones[4] values. An example of this wave-length error may be noted in Fig. 6. A somewhat similar error may possibly account for the deviation in the K–L region of copper. These obvious discrepancies do not appear for most elements.

It may appear to some that further elaboration of empirical expressions is unnecessary in view of developments in the classical theory of radiation, but as no simple method of calculating mass absorption coefficients accurately from classical theory has been made available, there is no alternative. Apparently there has been little interest in the measurement of absorption coefficients, however, for the only available observed data are old and incomplete. Published values often differ by 100 percent.

PREVIOUS EXPRESSIONS

Early empirical expressions assumed that the mass absorption coefficient for a given element was given by

$$(\mu/\rho)_Z = k\lambda^3. \tag{1}$$

Further investigation indicated that the mass absorption coefficient for a given element Z was roughly given by

$$\mu/\rho = KZ^4\lambda^3. \tag{2}$$

All observers have noted that these equations given approximate values only under suitable conditions. For high values of Z and λ they do not give even a good approximation.

Owen, in 1912[5] and 1918,[6] proposed Eq. (2) in its exact form, as did Bragg and Pierce.[7] The change in absorption with Z^4 was later verified by Richtmyer and Warburton.[8]

The variation of the absorption coefficient with λ^3 has been controversial, as may be expected, depending upon the wave-length under consideration. Richtmyer,[9] in 1925, found that his data could be represented by

$$(\mu/\rho)_Z = K\lambda^3 + b. \tag{3}$$

For the shortest wave-lengths for which K is valid this will be within the limits of experimental error, and b will represent the mass scattering coefficient. Deviations from the λ^3 law were, however, noted by Richtmyer,[10] Allen,[11] and others.[12,13]

An expression of form similar to Eq. (2) was devised theoretically from J. J. Thompson's old pulse theory[14] by Compton[15] in 1924.

Compton,[16] Kramers,[17] Oppenheimer,[18] and de Broglie[19] have suggested expressions based

TABLE II. Values assigned for Eq. (10) for wave-lengths less than λ_K.

Z	ν_1		ν_2		$3\nu_3$		
	n_1	n_2	n_1	n_2	n_1	n_2	s
1	1	2	1	3	1	∞	0
6	1	3	1	4	1	∞	½
13	1	3	1	4	1	∞	½
29	1	4	1	∞	1	∞	1
50	1	4	1	∞	1	∞	1
82	1	3	1	4	1	∞	1

[3] Trevor R. Cuykendall, Phys. Rev. **50**, 105 (1936).
[4] M. T. Jones, Phys. Rev. **50**, 110 (1936).
[5] E. A. Owen, Proc. Roy. Soc. **86**, 434 (1912).
[6] E. A. Owen, Proc. Roy. Soc. **94**, 522 (1918).
[7] Bragg and Pierce, Phil. Mag. **28**, 626 (1914).
[8] F. K. Richtmyer and F. W. Warburton, Phys. Rev. **22**, 539 (1923).
[9] F. K. Richtmyer, Phys. Rev. **27**, 1 (1925).
[10] F. K. Richtmyer, Phys. Rev. **18**, 13 (1921).
[11] S. J. M. Allen, Phys. Rev. **27**, 266 (1926).
[12] K. W. Wingardh, Zeits. f. Physik **8**, 363 (1922); *ibid.*, **20**, 315 (1923).
[13] E. Dershem and M. Schein, Phys. Rev. **37**, 1238 (1931).
[14] J. J. Thompson, *Conduction of Electricity through Gases* second edition, p. 326.
[15] A. H. Compton, Phys. Rev. **14**, 249 (1919).
[16] A. H. Compton, Phys. Rev. **14**, 249 (1919).
[17] H. A. Kramers, Phil. Mag. **46**, 836 (1923).
[18] J. R. Oppenheimer, Zeits. f. Physik **41**, 268 (1927).
[19] L. de Broglie, J. Physique **3**, 33 (1922).

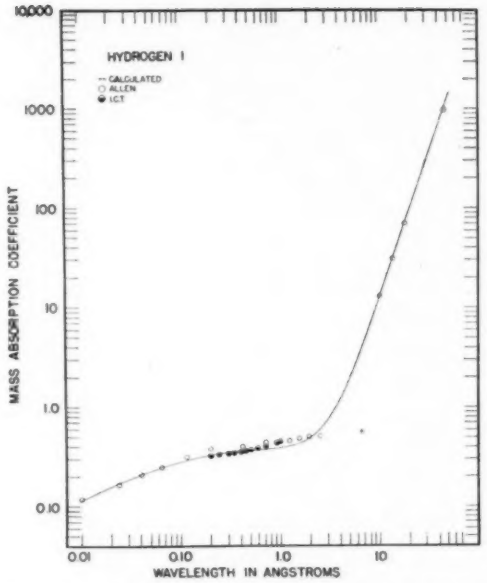


FIG. 1. Calculated and observed values of the mass absorption coefficient of hydrogen.

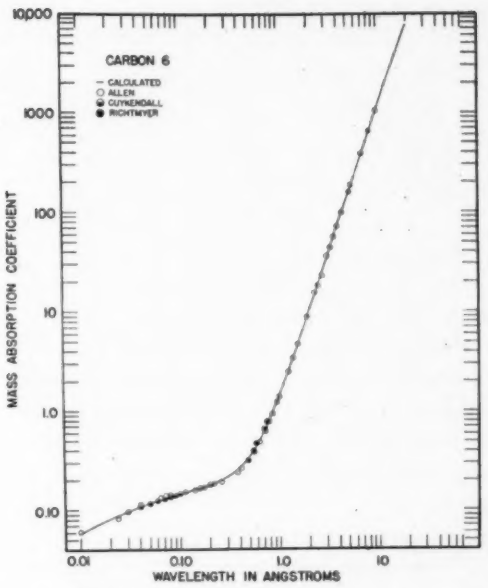


FIG. 2. Calculated and observed values of the mass absorption coefficient of carbon.

on theoretical considerations. Richtmyer,[20,21] Jonsson,[22] and Walter[23] have discussed general absorption formulas; Houston[24] and others[25] have commented on the failure of previous empirical formulas to represent adequately experimental data.

The author in a previous article[1] has shown that the photoelectric absorption coefficient τ_a of any one element is given with excellent accuracy by the expression

$$\tau_a = C/\nu^3 - D/\nu^4 = C'\lambda^3 - D'\lambda^4, \qquad (4)$$

where C and D are constants which change value between absorption discontinuities, and λ or ν is incident wave-length or frequency.

TABLE III. Values assigned for Eq. (10) for wave-lengths between λ_K and λ_L.

Z	ν_1		ν_2		$3\nu_3$		
	n_1	n_2	n_1	n_2	n_1	n_2	s
13	2	3	1	3	2	∞	4
29	2	3	1	3	2	∞	4
50	2	4	1	2	2	∞	4
82	2	5	1	2	2	∞	4

[20] F. K. Richtmyer, Phys. Rev. **27**, 1 (1925).
[21] F. K. Richtmyer, Phys. Rev. **30**, 755 (1927).
[22] E. Jonsson, Thesis, Uppsala (1928).
[23] B. Walter, Forts. a. d. **35**, 929, Geb. der Rontgen 1308 (1927).
[24] R. A. Houston, Proc. Roy. Soc. **40**, 35 (1920).
[25] A. H. Compton, *X-Rays and Electrons* (D. Van Nostrand Company, Inc., New York, 1926).

The constants C and D have been found to be given by

$$C = \nu_1\nu_2(3\pi e^2/mc), \qquad (5)$$

$$D = \nu_1\nu_2\nu_3(3\pi e^2/mc), \qquad (6)$$

where ν_1, ν_2, and ν_3 are critical frequencies obtained without justification from the accepted spectral emission formula

$$\nu_c = Rc(Z-s)^2\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)$$

$$+Rc\left[\alpha^2(Z-s)^4\left(\frac{1}{n_1^4} - \frac{1}{n_2^4}\right)\left(\frac{n}{k} - \frac{3}{4}\right) + \cdots\right]. \qquad (7)$$

Where R is Rydberg's number, c the velocity of light, Z is atomic number, α is Sommerfeld's fine structure constant, s is a screening constant, and n_1 and n_2 are quantum numbers.

Satisfactory agreement is obtained by assigning the following values: $S=0$, $\frac{1}{2}$, 1, etc., $n=1, 2, 3, 4$, etc., $\alpha=(2\pi e^2/ch)$, $k=n$, so $n/k=1$.

It is essential that the bracketed term of Eq. (7), which represents the Sommerfeld relativity correction, be included to sufficiently high powers of α^2 that the last additive term may be neglected. The values of C and D for any element can then be obtained by inserting the proper atomic number, Z, and assigning suitable values for s and n. The constants s and n are consistent,

TABLE IV. For wave-lengths in angstroms shorter than λ_K.

Z	C	D	Z/A	λ_K
1	1.274×10^{-2}	4.661×10^{-6}	0.99206	.....
6	1.2245	1.421×10^{-2}	0.49958	43.5
13	14.60	0.8362	0.48201	7.955
29	178.5	51.73	0.45619	1.377
50	940.2	853.8	0.42122	0.4239
82	4158	11040	0.39573	0.1407

so that it is possible to assign values to an element of atomic number Z which lies between two elements Z_1 and Z_2 for which values of s and n are satisfactorily known. When C and D are satisfactorily known for a given element for any one wave-length between discontinuities, all other wave-lengths in that region can be calculated.

Substituting from Eqs. (5) and (6) for C and D in Eq. (4) and multiplying by N_0/A, we have

$$\frac{\tau}{\rho}=\left(\frac{\nu_1\nu_2}{\nu^3}-\frac{\nu_1\nu_2\nu_3}{\nu^4}\right)\frac{3\pi e^2}{mc}\frac{N_0}{A}. \tag{8}$$

From the Klein-Nishina formula

$$\frac{\sigma}{\rho}=\frac{3}{4}\cdot\frac{8}{3}\frac{\pi e^4}{m^2c^4}\cdot\frac{N_0}{A}\cdot Z\left[\frac{1+\alpha}{\alpha^2}\left(\frac{2[1+\alpha]}{1+2\alpha}-\frac{1}{\alpha}\log_e(1+2\alpha)\right)+\frac{1}{2\alpha}\log_e(1+2\alpha)-\frac{1+3\alpha}{(1+2\alpha)^2}\right], \tag{9}$$

adding the right hand terms from Eqs. (8) and (9) and dividing and multiplying by $\frac{3}{4}[(8/3)(\pi e^4/m^2c^4)]$, we obtain

$$\frac{\mu}{\rho}=\left\{\left(\frac{\nu_1\nu_2}{\nu^3}-\frac{\nu_1\nu_2\nu_3}{\nu^4}\right)2\pi\left(\frac{3}{4}\frac{mc^3}{\pi e^2}\right)+Z\left[\frac{1+\alpha}{\alpha^2}\left(\frac{2[1+\alpha]}{1+2\alpha}-\frac{1}{\alpha}\log_e(1+2\alpha)\right)+\frac{1}{2\alpha}\log_e(1+2\alpha)-\frac{1+3\alpha}{(1+2\alpha)^2}\right]\right\}\frac{3}{4}\left(\frac{8}{3}\frac{\pi e^4}{m^2c^4}\right)\frac{N_0}{A}, \tag{10}$$

where $\alpha=h\nu/mc^2$.

The term $2\pi[(3/4)(mc^3/\pi e^2)]$ has the dimension of a limiting angular frequency, ω. The term $[(8/3)(\pi e^4/m^2c^4)]=\sigma$, the classical scattering per electron.

It may be noted that the large bracketed part of Eq. (10) is now dimensionless, and the mass photoelectric coefficient may be represented by

$$\frac{\tau}{\rho}=\left\{\left(\frac{\nu_1\nu_2}{\nu^3}-\frac{\nu_1\nu_2\nu_3}{\nu^4}\right)\omega\right\}\frac{3}{4}\frac{N_0}{\sigma}\frac{}{A}. \tag{11}$$

Values of μ/ρ calculated by means of Eq. (10) for frequencies beyond a critical absorption frequency are indicated by the dotted line in Fig. 4 for copper.

TABLE V. For wave-lengths in angstroms between λ_K and λ_L.

Z	C	D	Z/A	λ_L
13	0.5796	4.294×10^{-3}	0.48201	....
29	14.79	0.8456	0.45619	11.8
50	106.9	20.87	0.42122	2.75
82	621.2	352.8	0.39573	0.78

TABLE VI. Values of $\sigma_e N_0$—wave-length in angstroms maximum value = 0.4011.

λ	$\sigma_e N_0$	λ	$\sigma_e N_0$	λ	$\sigma_e N_0$	λ	$\sigma_e N_0$
0.0010	0.0262	0.010	0.114	0.10	0.282	1.0	0.384
0.0012	0.0300	0.012	0.125	0.12	0.295	1.2	0.388
0.0015	0.0350	0.015	0.140	0.15	0.310	1.5	0.392
0.0020	0.0431	0.020	0.159	0.20	0.329	2.0	0.394
0.0025	0.0502	0.025	0.175	0.25	0.339	2.5	0.396
0.0030	0.0566	0.030	0.188	0.30	0.352	3.0	0.397
0.0040	0.0680	0.040	0.210	0.40	0.362	4.0	0.397
0.0050	0.0778	0.050	0.227	0.50	0.367	5.0	0.398
0.0060	0.0866	0.060	0.243	0.60	0.374	6.0	0.399
0.0080	0.101	0.080	0.265	0.80	0.382	8.0	0.400

SUMMARY

The data of Chao[26,27] and Lauritsen[28] indicate that Eq. (10) gives satisfactory agreement with experiment on wave-lengths as short as 0.005A. But at very short wave-lengths additional types of absorption occur as a result of pair production, etc.

Equation (10) gives values of μ/ρ which are, on the average, within several percent of the values given in the original paper; and where differences exist, Eq. (10) appears to favor the observer's values.

Constants for the determination of the mass photoelectric absorption coefficient τ/ρ by Eq. (8) have been determined at long wave-length where other factors are apparently negligible. The mass scattering coefficient, σ/ρ, is determined by the Klien-Nishina formula, Eq. (9), which becomes the determining factor at short wave-lengths.

Usually the total mass scattering coefficient,

[26] C. Y. Chao, Phys. Rev. **36**, 1519 (1930).
[27] C. Y. Chao, Science Reports, National Tsing Hua University **1**, 59 (1932).
[28] J. Reed and C. C. Lauritsen, Phys. Rev. **45**, 433 (1934).

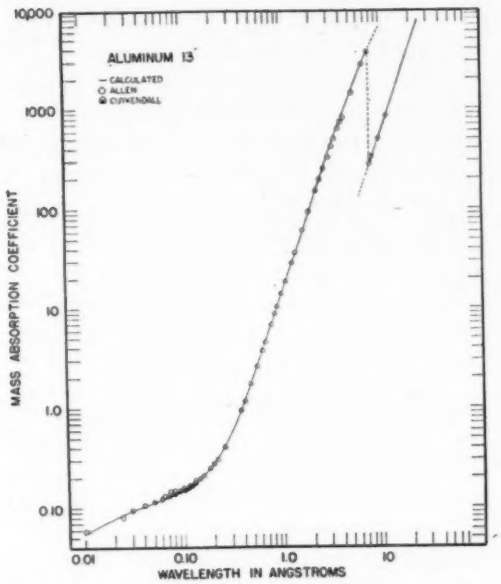


FIG. 3. Calculated and observed values of the mass absorption coefficient of aluminum.

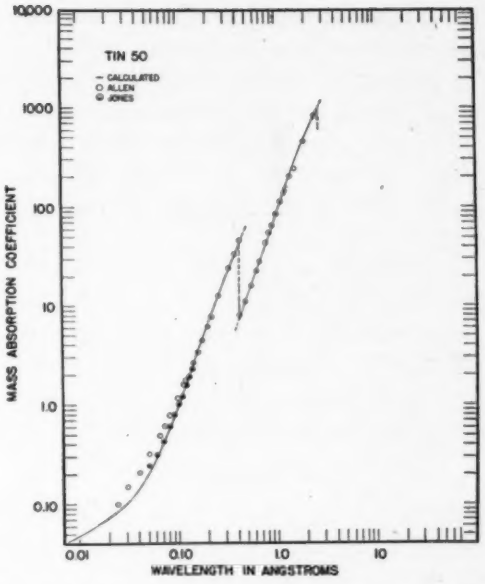


FIG. 5. Calculated and observed values of the mass absorption coefficient of tin.

σ/ρ, is considered as being the sum of σ_s/ρ and σ_a/ρ, where the latter represents the true absorption associated with scattering. It appears that when σ, the Klien-Nishina values of σ_s, are used, the addition of a term σ_a leads to values of μ/ρ which are too high, when compared with experimental values over a number of elements. We may therefore consider that for practical purposes from Eq. (10):

$$\frac{\tau}{\rho}+\frac{\sigma}{\rho}=\frac{\tau_0}{\rho}+\frac{\sigma_a}{\rho}+\frac{\sigma_s}{\rho}.$$

The values of μ/ρ given for several elements have not been selected as good examples of agreement with observed values. No obvious deviations are observed on most elements. Values

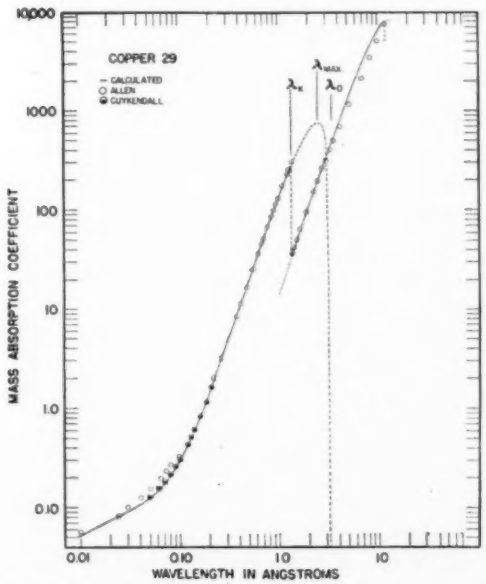


FIG. 4. Calculated and observed values of the mass absorption coefficient of copper.

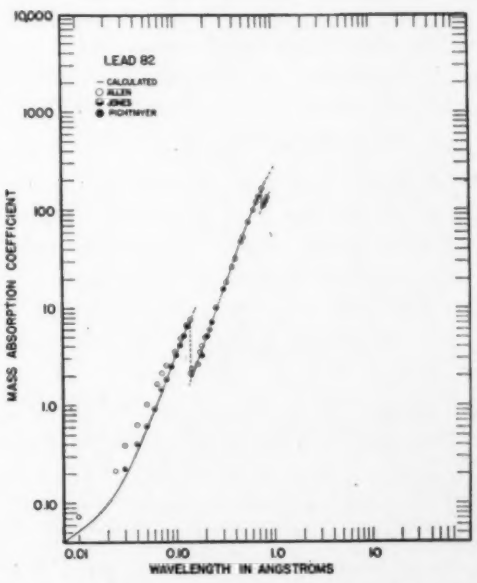


FIG. 6. Calculated and observed values of the mass absorption coefficient of lead.

of μ/ρ for other elements are being computed and will be presented later in tabular form.

The values of mass absorption coefficients previously published by the author are evidently a good representation of the probable values for practical purposes, but it is believed that the values obtained by the expressions given here are to be preferred.

All calculations were carried to five significant figures, although only four are shown in the data.

The author wishes to express his gratitude to Mr. William W. Managan for his generous effort in expanding Sommerfeld's original relativity correction[29] to the higher powers of α^2.

CALCULATIONS

Table I gives the values of μ/ρ for six elements calculated from Eq. (10), using the following values of the physical constants, taken from Birge:[30] N_0, 6.0228×10^{23}, mole^{-1}, h, 6.624×10^{-27}, erg sec., e, 4.8025×10^{-10}, abs e.s.u., c, 2.9977×10^{10}, cm sec.$^{-1}$, m, 9.1066×10^{-28}, g, R, 1.09737×10^{5}, cm^{-1}, α, 7.2976×10^{-3}.

Table II shows the quantum numbers assigned for the several elements for wave-lengths shorter than the K critical absorption wave-length.

Table III shows the values assigned for wave-lengths between λ_K and λ_L.

The values of the quantum numbers given in Tables II and III have been assigned arbitrarily. k is taken as always equal to n.

The data from Table 1, computed from the values in Table IV and V, are compared with observed values in Figs. 1–6.

Tables IV and V give values to be used in Eq. (3). In this convenient form, μ/ρ may be calculated for wave-lengths other than given in Table I. Table VI gives the values of $\sigma_e N_0$.

[29] A. Sommerfeld, *Atomic Structure and Spectral Lines* (Methuen and Company, Ltd., London), third edition, Vol. 1, p. 257.

[30] R. T. Birge, Rev. Mod. Phys. **13**, 233 (1941).

A Solution of the Problem of Rapid Scanning for Radar Antennae*

R. F. Rinehart
Case Institute of Technology, Cleveland, Ohio

(Received April 30, 1948)

The problem of devising a rapid scanning focusing antenna based on wave-guide principles has been solved only approximately and not too satisfactorily. In the present paper, a known property of a plane optical system with variable index of refraction is shown to satisfy the requirements of such an antenna system. By a suitable transformation, the plane variable refractive index system can be transformed into a surface of revolution of constant refractive index which reproduces the optical properties of the plane system. This surface provides an apparently practicable solution to the scanning problem. It is shown that the diameter of the feed circle can be made arbitrarily small by judicious use of an ellipsoidal reflector.

INTRODUCTION

THE problem of designing a rapid scanning radar antenna employing the theory of parallel plate wave guides has been discussed by S. B. Myers.[1] The fundamental property of such a curved wave guide is a Fermat least time principle: *Electromagnetic energy is propagated along geodesics* (i.e., *shortest distance curves*) *of the hypothetical mean surface between the parallel conducting plates.* The rapid scanning problem for microwaves, as educed in Myers' paper, is to find a surface Σ (to play the role of mean surface of a parallel plate wave guide) such that

(a) Σ contains a circle (the *feed circle*),

(b) Σ contains a straight line segment (the *aperture*),

(c) (Optical requirement) all geodesics joining a fixed point of feed circle to points of the aperture meet the aperture at a fixed angle,

(d) (Illumination requirement) all geodesics normal to the feed circle pass through a fixed point (the *center of illumination*) of the aperture.

Myers points out that it is not known whether a surface with such properties exists. If, however, requirement (d) is relinquished, a solution to the modified problem is the so-called RCA solution,

* This paper is a result of work performed on Contract No. W28-099-ac-141 for the United States Air Forces under the sponsorship of the Electronic Research Laboratory, Air Materiel Command. The particular results treated here grew out of a suggestion of Dr. Wade Ellis of the Watson Laboratories.

[1] S. B. Myers, "Parallel plate optics for rapid scanning," J. App. Phys. **18**, 221 (1947).

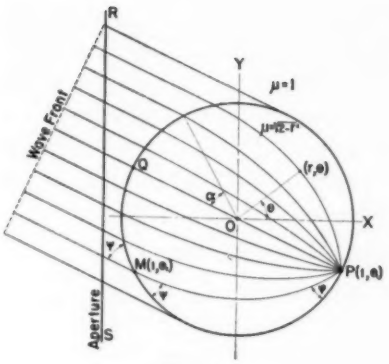


FIG. 1.

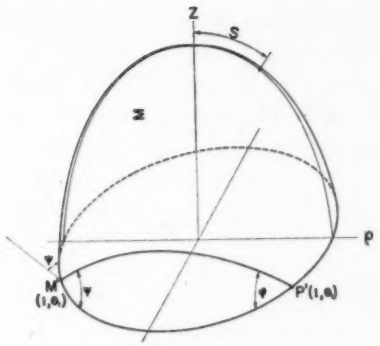


FIG. 2.

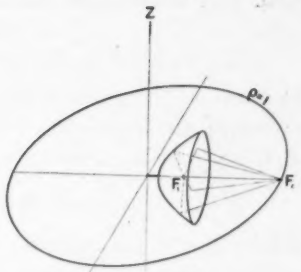


FIG. 3. Use of elliptical reflector to reduce feed circle diameter.

constructed from a portion of a frustum of an appropriate cone.[1] This solution furnishes unsatisfactory illumination characteristics at the aperture.

The purpose of this paper is to describe a solution surface satisfying requirements (a), (b), and (c) and which yields what seems to be satisfactory illumination.

A PERTINENT RESULT FROM OPTICS

The solution stems from an appropriate adaptation, or transformation, of the following known optical result:[2]

In the optical system composed of a unit circle with variable index of refraction $\mu=(2-r^2)^{\frac{1}{2}}$, where r is the distance from the center, imbedded in a plane medium of refractive index 1, all light rays entering the circle from any fixed point P on its periphery emerge from the circle parallel to the diameter OP[3] (see Fig. 1). Imagine for the moment that this unit circle is a mean surface for a variable refractive index wave guide. The unit circle would serve as feed circle. Any straight segment in the plane, not intersecting the circle, could serve as aperture with respect to the optical requirement. As P rotates through a semicircle, the emergent beam scans 180°. It may be remarked that this system also satisfies the illumination requirement with respect to a diameter of the circle as *virtual* aperture.

It can be shown that the angular width of a beam from P is equal to the angle subtended at 0 by the emergent beam. From this it follows that the illumination intensity along a cross section of the emergent beam varies as $\sec\alpha$. Thus a 90° beam width at P, for instance, would provide an emergent beam of linear width $\sqrt{2}$, and the illumination intensity would vary only 41 percent from its value at Q. This would seem to be an acceptable illumination pattern.

Such a system has the obvious disadvantage that it can utilize only a portion of each revolution (unless two opposite reflectors are used). On the other hand, this permits the feed to be active for an arc of only 90° in order to get a complete 90° scan by the emergent beam.

THE TRANSFORMATION TO A SYSTEM OF CONSTANT REFRACTIVE INDEX

We seek now a surface which duplicates the optical characteristics of the variable refractive index (v.r.i.) system just described. The straight wave front RS in the v.r.i. system results from the selective time delays induced in the various rays by the variable refractive index, which slows down the inner rays traveling the shorter distances. We seek to accomplish this delay by varying the path lengths, rather than the velocities, of the rays. We propose to achieve this effect by "bellying up" the unit circle into an appropriate surface of revolution.

To this end we seek a surface Σ generated by revolving about the z axis a curve $\rho=\rho(s)$, ρ being the arc length measured from the axis, which will reproduce the optical properties of the v.r.i. system (see Fig. 2). In the v.r.i. system, the differential of time (in appropriate units) along a curve is given by

$$dt^2=\mu^2 d\zeta^2=\mu^2(dr^2+r^2d\theta^2), \qquad (1)$$

[2] See, for instance, R. K. Luneberg, *Mathematical Theory of Optics* (Brown University Lecture Notes, 1944), pp. 212–213.

[3] Note that μ is continuous at $r=1$, and hence that incident and excident angles are equal at the unit circle.

where $d\zeta$ is the differential of arc length in the plane of Fig. 1, and (r,θ) are polar coordinates.

Along a curve of the surface Σ, the differential of time, in the same units, is precisely the differential of arc length $d\xi$, since the refractive index is 1, and is given by

$$dt^2 = d\xi^2 = ds^2 + \rho^2 d\theta^2, \tag{2}$$

where z, ρ, θ are cylindrical coordinates.

We desire to so choose the function $\rho(s)$ that dt as given by (2) will be identical with dt as given by (1). Let us introduce a new (to Σ) variable r by means of the relation

$$ds = [\rho(s)/r]dr. \tag{3}$$

(2) then takes the form

$$d\xi^2 = \rho^2/r^2(dr^2 + r^2 d\theta^2) = (\rho^2/r^2)d\zeta^2 \tag{4}$$

which will be identical with (1) provided the relation

$$\rho/r = \mu(r) = (2-r^2)^{\frac{1}{2}} \tag{5}$$

is imposed. From (5) follows, since $0 \leqq r \leqq 1$,

$$r = [1+(1-\rho^2)^{\frac{1}{2}}]^{\frac{1}{2}}. \tag{6}$$

Relations (3) and (6) provide a differential equation in s and ρ, which, with the boundary condition, $s=0$ when $\rho=0$, defines uniquely the functional relation between s and ρ which makes $dt = d\xi$.

From Eqs. (5) and (6) we note that when $\rho=0$, $r=0$ and when $\rho=1$, $r=1$. Hence if we cut off the surface at $\rho=1$ to obtain our surface Σ, Eq. (5) defines a map of Σ on the unit circle in the (r,θ) plane of the v.r.i., in which

$$\rho \longrightarrow r \quad \text{and} \quad \theta \longrightarrow \theta.$$

Now any curve arc which minimizes $\int dt$ in the v.r.i. system will have an image on Σ which minimizes $\int d\xi$, and conversely. Hence light rays in the v.r.i. system have images which are geodesics of Σ, and conversely.

Furthermore, Eq. (4) shows that this map is conformal (with magnification ratio ρ/r) and hence angles between curves are preserved under the mapping. Thus if a ray in the v.r.i. system leaves $P(1,\theta_1)$ at an angle φ with the unit circle, and recontacts the unit circle at $M(1,\theta_2)$ at an angle ψ with the unit circle, then the image geodesic of Σ behaves in precisely this same way with respect to the image circle $\rho=1$.

Now let the surface Σ be continued planewise beyond $\rho=1$. It is well known that a geodesic makes equal angles, approaching and leaving, with an edge of a surface. It follows that the surface with the planar continuation for $\rho=1$ duplicates completely the optical properties of the v.r.i. system, and constitutes a solution of the scanning problem.

PROPERTIES OF THE SURFACE

The equation of the surface Σ is readily computed from (3) and (6) to be

$$s = 1/2(\rho + \sin^{-1}\rho). \tag{7}$$

Its equation in cylindrical coordinates can be computed in terms of elementary functions. It is

$$\begin{aligned} z = 1/2[&4-3\rho^2+4(1-\rho^2)^{\frac{1}{2}}]^{\frac{1}{2}} \\ &-1/\sqrt{3}\,\log([3/2\{1+(1-\rho^2)^{\frac{1}{2}}\}]^{\frac{1}{2}} \\ &+[1/2\{3(1-\rho^2)^{\frac{1}{2}}+1\}]^{\frac{1}{2}}) \\ &-1/2+1/2\,\sqrt{3}\,\log(2+\sqrt{3}). \end{aligned} \tag{8}$$

The surface has a horizontal tangent plane at $\rho=0$ and a vertical tangent plane at $\rho=1$. The surface is concave down everywhere between $\rho=0$ and $\rho=1$.

REDUCTION OF SIZE OF FEED CIRCLE

The parallel plate system described here seems practicable from every standpoint except the large size of the feed circle, relative to the emergent beam width, or effective aperture. It can be proved that this difficulty is geometrically insurmountable in any surface of revolution. Specifically, the linear width of the emergent beam can never exceed the diameter of the feed circle (or of any intervening parallel circle of the surface). Consequently, in the solution given here the feed circle is the smallest attainable. It is possible, however, to reduce the size of the feed circle by optical means. A reflector in the shape of a section of an ellipsoid can be utilized to produce a *virtual* point source on the unit circle from an actual point source situated much closer, or even *on* the axis of the surface. As the actual feed F_1 (Fig. 3) rotates on a small circle, the virtual, or image, feed F_2 describes the unit circle. Rough computations indicate that a reflector yielding a suitable angular beam width could be designed to be housed beneath the dome-shaped wave guide. With this refinement the antenna system described seems to furnish a workable solution to the scanning problem.

Letters to the Editor

Electron Optical Observation of Magnetic Fields

L. MARTON
National Bureau of Standards, Washington 25, D.C.
June 21, 1948

IN a previous communication[1] it was pointed out that the electron optical "Schlieren" effect can be used for the observation of minute magnetic and electrostatic fields. Reference was also made to the use of that method for the purpose of quantitative measurement of the fringe field, extending from magnetized regions of a steel wire, as used for magnetic recording purposes. In pursuing observations on the same subject we used a center stop catching all direct radiation and allowing only the deflected electrons to reach the fluorescent screen. This center stop consisted of a 1-mm diameter thin copper disk supported by two silver cross-wires of 0.033-mm diameter. In one of our experiments it so happened that the image of the center stop (produced by projection from the virtual source between the lens and the stop) was displaced laterally by means of an external deflecting magnetic field. By adjusting the image of one of the supporting silver wires to be parallel with the image of the magnetized wire, we found that the image of the silver wire is not a straight line when in the neighborhood of the image of the magnetized wire. Figures 1–3 show a sequence in which the two images are first completely separated, then moved closer, and finally made to coincide. It is found that in the case of the two images, the image of the thin silver wire bulges out everywhere where magnetized regions of the steel wire deflect the incoming electron beam.

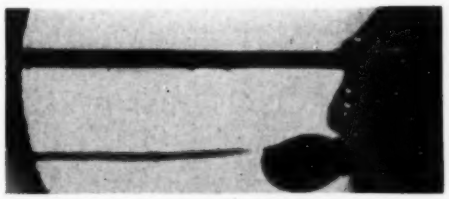


FIG. 1.

FIG. 2

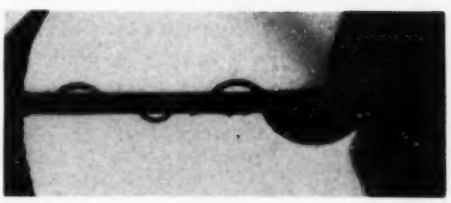


FIG. 3.

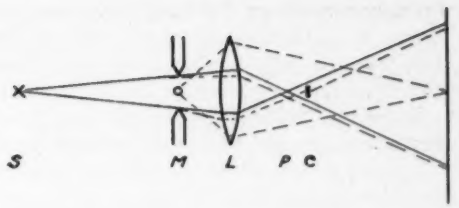


FIG. 4.

A schematic explanation of this observation is given by Fig. 4. The magnetic lens L used for forming a somewhat magnified image of the magnetized wire M forms a reduced image of the distant source S at the point P. This reduced image of the source serves as a virtual source forming a shadow image of the center stop C and of its supporting wires. If the wire M has magnetized regions, however, the electron beam coming from the source will be deflected at each such magnetized region with the result that the virtual source P will be displaced sidewise. Such a displacement produces the images as observed, and this displacement is obviously a function of the strength of the magnetic field.

In close analogy to the electron optical "Schlieren" effect, the displacement of that image can be used for a quantitative measurement of the magnetic field strength. For this purpose it is assumed, as before, that the magnetic field of the wire is produced by an ideal dipole and thus the field obeys the inverse cube law. The absolute value of the field can be calculated at least to a first approximation at the equatorial plane of the dipole from the measured deflection of the image and from the known geometrical constants of the apparatus. Calculations by Mr. S. H. Lachenbruch (to be published later) indicate that, for a commercially available magnetic recording wire, a recording, produced by 40-ma exciting current of the recording head, produces a magnetic field of 100 gauss at 0.1-mm distance from the wire axis. These data are based on measurement carried out on Fig. 5.

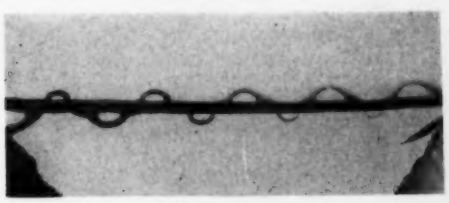


FIG. 5.

It is to be pointed out that this method offers, in some cases at least, the distinct advantage over the "Schlieren" observation that the displacement of the wire image can be measured with greater accuracy than the intensity distribution across the "Schlieren" pattern.

An extension of this method consists of placing a network of cross-wires beyond the back focal plane of the imaging lens (at the position *C* of Fig. 4). Such a network enables us to record a complete mapping of extended fields in a single operation. Experiments are being undertaken here to apply this method to numerous problems.

[1] L. Marton, J. App. Phys. **19**, 687 (1948).

An Observation of Corrosion Cracking without Stress

W. D. Robertson and H. H. Uhlig
Department of Metallurgy, Massachusetts Institute of Technology, Cambridge, Massachusetts
June 21, 1948

STRESS corrosion cracking of metals and alloys is one of the most important unsolved problems in the field of corrosion. The reasons are largely unknown why metals subjected to residual or applied stress suddenly crack after a period of time in specific media. Various suggestions of mechanism have been made,[1-3] all pointing to the presence of metal or non-metal acting as anode at the base of a growing crack with the wall functioning as cathode. However, the detailed mechanism is not entirely clear, particularly with respect to transcrystalline cracking and to the extreme selectivity of the environment in which cracking occurs. The necessity of a specific environment for cracking indicates that the explanation based alone on galvanic effects of stressed metal may be inadequate and too general since, presumably, many conducting environments should thereby cause stress corrosion cracking. Such is not the case. Furthermore, non-metals such as plastics exhibit stress cracking phenomena in various organic solvents where an electrochemical mechanism is obviously irrelevant.

Refined experiments on stress corrosion cracking or "season cracking" of brass, although not pointing to a specific mechanism, have indicated the relative importance of some factors.[4] Thus, impurities do not play a large role, and the addition of other elements does not greatly reduce susceptibility; furthermore, the critical stress necessary for cracking in an ammonia atmosphere has been reduced to lower and lower limits. Single crystals of brass in ammonia also crack by stress corrosion, so the grain boundary in polycrystalline alloys is relieved of total responsibility for failure. One is tempted to conclude from these facts that corrosion cracking in some alloys may be

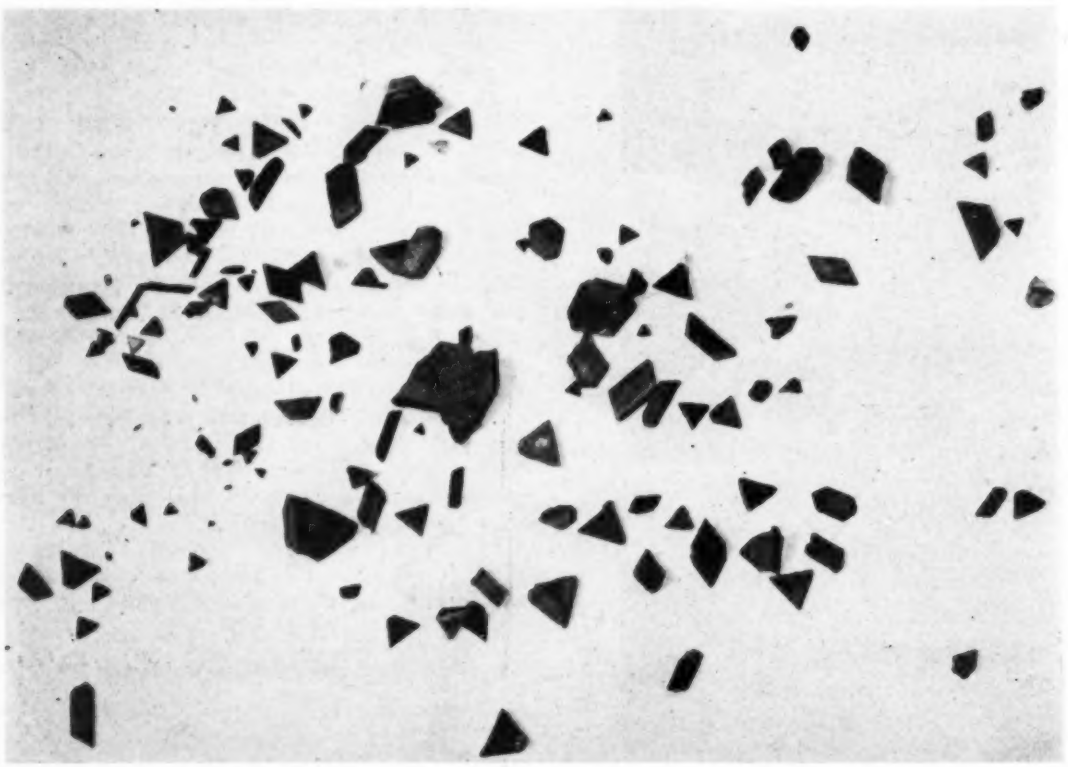


Fig. 1. Crystallographic disintegration of Mg_2Sn single crystal.

associated with metallurgical factors inherent to the alloy itself.

An observation we recently made, in connection with a study of the chemical properties of the intermetallic compound Mg_2Sn, lends weight to this possibility. The compound was prepared in an argon atmosphere from sublimed magnesium (courtesy of the Dow Chemical Company) and from electrolytically purified tin. Large crystals were produced from the liquid by cooling slowly through the melting point to room temperature at a rate of 20°C per minute. The resulting cast rods were $\frac{3}{8}$" in diameter and could be fractured (cleaved) easily to produce a bright bluish colored, essentially plane cleavage face across the entire diameter of the rod.

These crystals react readily with oxygen-free sodium chloride solutions to form magnesium hydroxide and finely divided tin or, in the presence of oxygen, stannous oxide. In water, however, the compound is comparatively stable, and in a moist atmosphere (30 percent relative humidity) the cleavage planes remain bright for a period of weeks.

Immersed in ordinary or triple distilled water for several hours, single crystals of the compound were observed to crack along definite planes, forming small geometrical fragments. Under the low power microscope, the fragments resolve themselves into three-sided pyramids, rhombohedra, and equilateral triangles, all of which can be derived from the octahedral cleavage planes of the crystal system involved. A typical group of crystal fragments, obtained by the action of distilled water on a single crystal of Mg_2Sn, is shown in Fig. 1. Similar crystallographic breakdown occurs in dilute sodium hydroxide, at a rate increasing with the hydroxyl ion concentration, in hydrogen peroxide (35 percent), and in water from which dissolved oxygen has been removed by boiling. Furthermore, the fragments continue to subdivide if allowed to remain in contact with water.

The crystal structure of Mg_2Sn is that of CaF_2 with the magnesium ions occupying the fluorine positions, and tin ions the positions of the calcium ions. The {111} planes are the cleavage planes across which cohesion presumably is least and are the planes of dense packed magnesium and tin ions, with the magnesium ions slightly elevated above the plane of tin ions. The action of water in splitting the crystal, although possibly related to traces of impurities in the compound, is more likely one of reaction or hydration of ions in the {111} planes, thereby destroying the bonds between these planes. Finite fragments are obtained, possibly owing to the fact that the crystal is not perfect, and water, therefore, is drawn in between certain preferred planes rather than into all planes equally. Allied with this mechanism, the formation of a corrosion product between cleavage planes may result in a pressure that mechanically splits the crystal.

The presence of hydrogen as a corrosion product does not appear to be involved, insofar as a crystal mounted as cathode in 0.02*N* NaOH (approximately 0.05 amp./cm²) fractured much less rapidly than another crystal, mounted in the same cell, as anode.

The slow cooling of the Mg_2Sn precludes large residual stresses, and the brittle character of the compound is such that large stresses could not be supported. Therefore, the fracturing appears to be an instance of transcrystalline corrosion cracking in the absence of appreciable stress.

The analogy with cracking in certain stressed magnesium or aluminum alloys in presence of water is obvious. The mechanism may have its counterpart in other metallic systems, with hydration and solvation supplementing galvanic action within the crack. Once a crack starts, application of tensile stress accelerates propagation of the crack and eventual failure of the alloy.

The authors are indebted to the Office of Naval Research for support of this investigation.

[1] E. H. Dix, Jr., Trans. A.I.M.E. **137**, 11 (1940).
[2] J. T. Waber, H. J. McDonald, and B. Longtin, Trans. Am. Electrochem. Soc. **87**, 209 (1945).
[3] R. B. Mears, R. H. Brown, and E. H. Dix, Jr., "Symposium on stress corrosion cracking of metals," A.S.T.M. and A.I.M.E. ,Philadelphia, 323 (1945).
[4] C. L. Bulow and G. Edmunds, "Symposium on stress corrosion cracking of metals," A.I.M.E. and A.S.T.M., Philadelphia, 19 and 67 (1945).

New Booklets

Tracerlab, Inc., 55 Oliver Street, Boston 10, Massachusetts, has issued its Catalog A, having 40 pages. Available on request.

James G. Biddle Company, 1316 Arch Street, Philadelphia 7, Pennsylvania, has announced a new bulletin, No. 24–25, on its "Ducter" low resistance ohmmeter, a handbook on low resistance measurements in the shop and in the field. 12 pages. Available on request.

Eberbach and Son Company of Ann Arbor, Michigan, offers free copies of its latest *Announcer of Scientific Equipment*, No. 48–6–31. The issue features an article on "Development of chemical porcelain."

Burrell Technical Supply Company, 1942 Fifth Avenue, Pittsburgh 19, Pennsylvania, has published bulletin No. 214 describing combustion tubes and boats and featuring a new tube, the "Zircum" combustion tube. 4 pages.

T. C. Wheaton Company, Millville, New Jersey, has published a bulletin illustrating and describing its line of electronic glassware. 4 pages.

Allen B. Du Mont Laboratories, Inc., 1000 Main Avenue, Clifton, New Jersey, announces the fifth edition of its *Catalog of Equipment for Oscillography*. 176 pages.

Ward Leonard Electric Company, Mount Vernon, New York, has issued its Catalog D-130 on resistors, rheostats, and radio amateur relays. 24 pages. A copy may be procured by writing to Radio and Electronic Distributor Division, Ward Leonard Electric Company, 53 West Jackson Boulevard, Chicago 4, Illinois.

National Bureau of Standards has prepared and made available as NBS Applied Mathematics Series 2 a 20-page *Table of Coefficients for Obtaining the First Derivative without Differences*. May be obtained only from the Superintendent of Documents, U. S. Government Printing Office, Washington 25, D. C., at 15 cents per copy.

The Scientific Instrument Manufacturers' Association of Great Britain, Ltd., 26 Russell Square, London W. C. 1, has published the 1948 edition of its handbook of electronic instruments, which gives a brief review of the products of members of the Electronics Section of the Association. Copies available upon request. 28 pages.

Here and There

New Appointments

Haldon A. Leedy, Acting Director of Armour Research Foundation of Illinois Institute of Technology, has been named Director.

William J. Youden, mathematical statistician and chemist, has been appointed to the staff of the National Bureau of Standards as Assistant Chief of the Statistical Engineering Section.

L. B. Spinney Honored

Professor L. B. Spinney, physics textbook author and head of the department of physics at Iowa State College, was honored recently by the Iowa State College Alumni Association when he was awarded a faculty citation in recognition of long, outstanding, and inspiring service to the college. Professor Spinney began his teaching duties at Iowa State in 1893.

Necrology

E. F. Burton, Chairman of the Department of Physics of the University of Toronto, passed away on July 6. He was outstanding for his pioneer work in electron microscopy on this continent.

Professor Emeritus Ernest George Merritt, first dean of the Graduate School at Cornell University and a member of the Cornell Physics Department for 45 years until his retirement in 1935, died June 5 at the age of 83.

Bulletin of Mathematical Biophysics

Following is the table of contents for the September 1948 issue of the *Bulletin of Mathematical Biophysics*:

"Physical aspects of organic evolution"—Alfred J. Lotka

"Theory of the measurement of blood flow by the dilution of an indicator"—John L. Stephenson

"Interaction of cycles"—J. B. Roberts

"An analysis of theoretical systems of differentiating nervous tissue"—Alfonso Shimbel

"Cycle distributions in random nets"—Anatol Rapoport

"On periodicities in metabolizing systems"—N. Rashevsky

"The kinetics of blood coagulation"—John Z. Hearon

"A note on the units of membrane permeability to water"—H. D. Landahl

"On the theory of blood-tissue exchange of inert gases: VI. Validity of approximate uptake expressions"—Manual F. Morales and Robert E Smith

The University of Chicago Press, Chicago 37, Illinois, Volume 10, Number 3.

Conference at Oak Ridge

The physics division of Oak Ridge National Laboratory sponsored a conference on "Nuclear Physics and Low Temperatures" the weekend of August 7 and 8. The conference, which was the first of a projected series of physics meetings was organized in response to the increasing interest in nuclear problems involving low temperature phenomena. Speakers and their subjects were as follows:

Alignment of nuclear spins—M. E. Rose, Oak Ridge National Laboratory.
Nuclear relaxation phenomena at low temperatures—E. M. Purcell, Harvard University.
Thermal equilibrium at low temperatures—H. B. G. Casimir, Philips' Works, Eindhoven, Netherlands.
The superfluid state—F. London, Duke University.

High Polymer Meetings at Brooklyn

The following are among future meetings on high polymers to be sponsored by Polytechnic Institute of Brooklyn, 99 Livingston Street, Brooklyn 2, New York: Symposium on Ultrasonics and Macromolecules, October 23, 1948, Brooklyn; Society of Rheology meeting, probably November 5–7, 1948, New York.

U. S. Testing Company Expands

In entering the field of supersonic research, the United States Testing Company, Hoboken, New Jersey, recently announced the completion of arrangements to take over in specific fields the research and development activities of the Crystal Research Laboratories of Hartford, Connecticut.